bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Utilizing a resource of enrichment profiles in plasma for the systematic assessment of antibody selectivity.

Claudia Fredolini [ǂ], Sanna Byström [ǂÞ], Laura Sanchez-Rivera [ǂÞ], Marina Ioannou [ǂ], Davide Tamburro [¥], Rui M. Branca [¥], Peter Nilsson[ǂ], Janne Lehtiö [¥] and Jochen M. Schwenk [ǂ,*]

[ǂ]Affinity Proteomics, Science for Life Laboratory, School of Biotechnology, KTH - Royal Institute of Technology, Box 1031, SE-171 21 Solna, Sweden

[¥]Cancer Proteomics, Department of Oncology-Pathology, Science for Life Laboratory, Karolinska Institute, Box 1031, SE-171 21 Solna, Sweden

[Þ] These authors equally contributed to the work

*Corresponding author:

Jochen M Schwenk, Affinity Proteomics, Science for Life Laboratory, School of Biotechnology, KTH - Royal Institute of Technology, Box 1031, 171 21 Solna, Sweden,

E: jochen.schwenk@scilifelab.se; T: +46 (0)8 790 9869

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ABSTRACT

There is a strong need for procedures that enable context and application dependent validation of antibodies. Here we describe the foundation for a resource aiding more detailed assessment antibody selectivity for capturing endogenous proteins from human plasma. In 414 immunoprecipitation (IP) experiments with EDTA plasma, data was generated by mass spectrometry (LC-MS) with 157 antibodies (targeting 120 unique proteins). Out of a total of 1,313 unique proteins, 426 proteins (33%) were detected in > 20% of the assays and indicate a background comprised of mainly proteins from the complement system. For all proteins identified either in heat-treated or untreated EDTA plasma, frequencies of occurrence were derived. We determined z-scores for each IP as a measure of enrichment to annotate the antibodies into four categories (ON-target, CO-target, OFF-target and NO-target). For 45% (70/157) of the tested antibodies, the expected target proteins were enriched (z-score ≥3) above background. There were 84% (59/70) of binders that co-enriched other proteins beside the intended target, either due to OFF-target binding or predicted interactions. Comparing several antibodies raised against IGFBP2, the established library allowed us to describe protein complexes in plasma, and we employed multiplexed sandwich immunoassays to confirm these. In summary, the generated resource of plasma enrichment profiles and background proteins adds a very useful and yet lacking starting point for the assessment of antibody selectivity in this clinically important body fluid. The provided insights will contribute to a more informed use of validated affinity reagents and may lead to further advancements of plasma proteomics assays.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

MAIN TEXT

Antibodies are important tools used in a wide range of assays within life science, but there is a growing awareness about the importance to assess the quality of the data generated therewith [1]. To address this challenge, the recently formed International Working Group for Antibody Validation (IWGAV) proposed five strategies to assess the experimental performance of antibodies [2]. However, the opportunities for evaluating antibodies in the given context (= sample type) and application (= assays) are limited. In particular for body fluids there are currently no tools for modulating the samples to overexpress or delete expression of a target of interest. While the use of GWAS provides powerful opportunities to assess the relation between plasma protein abundance and genetic information [3, 4], affinity reagents are preferably still validated before the intended use. There is further need to experimentally assess selectivity of the affinity reagents and to enable the development of sensitive assays and technologies for proteins in plasma or serum.

Here, we aim to describe the foundation of a resource for the assessment of antibodies used to capture proteins from plasma. We utilize immunoprecipitation (IP) of full-length proteins in combination with mass spectrometry (MS) to systematically describe a library of identified proteins. In more than 400 IPs with plasma frequencies of occurrence (*f*) and enrichments scores (*z-scores*) were collected, and we list those proteins that were commonly detected as plasma background. Similar efforts have been applied to evaluate the performances of antibodies for IP in cell lysates [5]. Apart from few studies focused on specific, smaller number of targets [6, 7] however, there are no large scale and systematic studies extensively applying IP of endogenous full-length proteins and MS for antibody validation in plasma. Otherwise, peptide-specific antibodies to determine protein abundance after digestion of plasma proteins have been more frequently applied in combination with MS readout [8-10]. Additional approaches such as iMALDI [11] and MS-based immunoassays [12] complement activities using protein-enrichment before MS analysis. The composition of the plasma proteome was recently updated and now lists around 3,500 proteins detectable using MS techniques [13]. Adding further targets to this list, foster hence to perform deeper quantitative analysis of plasma proteins, and will require large recourses of validated antibodies. For MS-based techniques, indeed one of the keys is to use enrichment as a strategy to enhance the sensitivity [14], while purely affinity-based techniques will greatly benefit from the utility of highly selective binder in multiplexed assay systems [15].

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Study Overview

To establish a plasma-centric resource for selectivity analysis of antibodies, we applied a systematic approach using 157 antibodies (targeting 120 proteins) and a workflow depicted in **Fig. 1A.** The assays were built on a previously described procedure, in which antibodies are covalently coupled onto magnetic polystyrene beads prior to incubation with the sample [16]. The study included binders both of monoclonal and polyclonal origin and different species. In order to compare the performance of different antibodies raised against a common antigen, a subset of 25 proteins (21%) were targeted by more than one antibody (**Supplementary Fig. 1**). The selection of presented targets was driven by giving priority to antibodies raised against proteins known to be part of the plasma proteome and to be associated to a disease (**Supplementary Excel Table, sheet: "Selected Targets"**). As described in **Fig. 1B,** the majority of binders (65%, N=101) were raised against target proteins detected previously 'in plasma' (48% N=75) or annotated as 'extracellular' (17%, N=26), and fewer proteins were annotated as 'cellular' (36%, N=56). As reference for protein abundance in plasma, we considered the estimated concentrations found in the 2017 built of the plasma PeptideAtlas [13].

Pilot study

In a pilot, we compared the experimental conditions to reach an analytical sensitivity to detect plasma proteins and to assess the potential to discriminate between specifically captured or frequently observed proteins (denoted plasma background). A set of antibodies obtained from commercially available ELISA kits for plasma protein analysis were used to cover a plasma concentration range between µg/ml (CRP) to low pg/ml (IL1A, see **Supplementary Note 1**). We tested combinations of different amounts of plasma (30-1000 µl) as well as increasing amounts of beads (30,000 - 1,000,000) to assess the detectability of the targets **Supplementary Fig. 2A**. We concluded that 100 µl of plasma and 500,000 beads were suitable for larger scaled investigations, enabling the detection within µg/ml and high pg/mL (e.g. KLK3 was detected in a pool of plasma obtained by mixing samples from healthy females and males, see **Supplementary Fig.2B)**. For the detection of pg/ml plasma proteins, such as IL1A, the IP may require even larger amounts of sample and beads, which was found less suitable for large scale efforts **(Supplementary Fig. 2A)**.The LFQ intensities from the IPs were used to calculate z-scores and compare peptides from replicated assays. In the pilot, all the antibody profiles barely reached an enrichment threshold (z-score ≥ 3) when using only this limited set of assays (21) in the library for data analysis. Utilizing a larger library of plasma IP data, as introduced and described below, improved the number of expected target proteins with z-scores ≥ 3 **(Supplementary Fig 2B).**

Assessing the resource

It is well accepted that MS provides in-depth information about the protein content of a sample, however hundreds if not thousands of proteins can be identified in a single IP experiment [17]. This calls for a careful assessment and interpretation of data from IP assays where many other proteins than the intended target can be identified in the same range of spectral counts or precursor intensities. As

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

described in the context of cell lysates, the necessity to compare the outcome of several experiments, including negative controls or unrelated antibodies is essential [5, 18, 19]. Mellacheruvu and colleagues described how lists of background proteins that were obtained from control assays performed in similar experimental conditions can be used to assess specific enrichments in IPs [18]. In our case, all the data (not exclusively control IgG) was generated in separate experiments (refer to as "experimental batches" or "batches"). These sets of data were used to build a larger library of proteins and annotate these by their frequency of identification (*f*) and enrichment scores (z). These then serves as a resource to assess enrichment profiles of the antibodies. Details on experimental batches and frequencies for proteins are described in **Supplementary Excel Sheets "Experimental Batches"** and **"Frequencies of identification"**.

We applied MaxLFQ label free quantification to 414 plasma IPs (pIP), a total of 1,313 unique proteins were identified, excluding variable domains of immunoglobulin for heavy and light chains. The complete list of proteins can be found in **Supplementary Excel Sheet**: **"Frequencies of identification"**. Per IP, about 290 proteins were detected on average. Data was collected from experiments performed under heat treatment versus not heat conditions, as further discussed below, we found that the average number of identifications per experiment were slightly higher in heat-treated plasma (300±97 in 276 pIPs) compared to untreated plasma (283±97 in 138 pIPs). For both sample types, the majority of proteins (66%) were detected in f < 20% of all assays (**Fig 1C** and **Supplementary Excel Sheet: "Frequencies of identification"**). Combining the frequencies of how many times a protein occurred in each IP, proteins found in f > 20% of all assays (denoted 'frequent') were compared with those found in f < 20% (denoted as 'less frequent'). The resulting GO analysis revealed that terms related to the complement activation and wound healing (GO:0002576, GO:0006956, GO:0050817, GO:0009611, GO:0007596) were enriched for the frequent proteins (**Supplementary Excel Sheet: "GO enrichment analysis").** Other enriched terms for the frequent proteins were related to lipoprotein and their complexes (GO:1990777, GO:0032994, GO:0034358) as well as vesicles (GO:0031983, GO:0060205). Further investigations also found a significant association (p-value < $2*10^{-16}$) relation between the frequencies and estimated protein abundance in plasma [13], as well as to sequence coverage (**Fig 1D, Supplementary note 3** and **Supplementary Figure 3A-B**). Further details on frequency of identification and intensities variates along the experimental batches can be found in the **Supplementary Figure 4A-B** and **Supplementary Excel Sheet:"Batches Kruskal Wallis test"**.

Differences between untreated and heat-treated plasma

Previously, we have shown that after heat treatment of plasma samples at 56°C, improved the limit of detection in plasma profiling assays using antibody bead array [20]. Heat treatment of plasma may indeed retrieve epitopes and making proteins more accessible to antibody binding, however heat may

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

also affect the sample's composition by causing aggregation and proteins to precipitate. As shown for all IPs in **Fig. 1C and Supplementary Figure 4A**, heat treatment may influence which proteins are listed among the common contaminants and their frequencies. When we consider LFQ intensities instead of frequencies, we observed that in heat-treated samples particularly fibrinogens (FGA, FGB and FGG) were more abundant (**Supplementary Fig. 5**). Previous observations state that denaturation of fibrinogen starts at 55ºC (two transitions peaks at 57.7ºC and 96ºC) and particularly affects the D fragment [21]. In heat-treated samples, the affinity of fibrinogens to plastic surfaces has also been reported to increase [22]. Via a similar mechanism, fibrinogen's unspecific binding to surfaces of magnetic beads or to heavy chains of IgG antibodies on the beads may be enhanced [23, 24]. Hence, heat-induced denaturation of fibrinogen may alter the frequency with which proteins of this family are identified. Other plasma proteins such as albumin, apolipoproteins (APOB, APOC2, APO2, APOE), Keratines (KRT1, KRT2, KRT10), Fibulin, Fibronectin 1 or IgM were less frequent among the common contaminants (see also **Supplementary Excel Sheet: "Frequencies of identification"**).

Classification

Extending the use of the library containing common contaminants and frequencies, we aimed at classifying the antibodies based on their enrichment profiles in plasma. Considering a protein as being enriched when the LFQ intensity was z-score ≥ 3, all antibodies were annotated according to the following categories **(Fig. 3A)**:

- ON-target: when a $z \geq 3$ was only assigned to the expected target.
- CO-target: other proteins besides the expected targets were detected with $z \geq 3$. Here we will discuss three sub-categories related to the origin go the additional target.
- OFF-target: proteins other than the expected targets were enriched with $z \geq 3$.
- NO-target: all detected proteins were classified with $z < 3$.

Knowing that common plasma background proteins may differ between assays using untreated and heat-treated plasma, data from the respective assays was analyzed separately and the resulting z-scores were combined for the global assessment. For the 157 antibodies, we made 1173 identifications for 681 unique proteins with $z \geq 3$. In each of the 414 IPs, we detected about 550 proteins, of which 6-9 proteins were enriched above the threshold. The combined outcome of the analyses is shown in **Table 1**, where the classification of 70 out of 157 antibodies (45%) was found supportive by falling into the ON- or CO-target category. When assessing only those antibodies targeting proteins previously annotated in plasma, the fraction increased to 61% (46/75). It is noteworthy that annotating antibodies in non-supportive categories could be due to limited affinity of the antibody, limited assay sensitivity or absence of the target in the used pools of plasma derived from healthy donors, as well as limited peptide detectability.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ON-target enrichment

Applying IP-MS analysis has been reported to improve the sensitivity of protein quantification [8, 12, 25]. Hence, IP-MS may also be used to detect lower abundant proteins, adding to those that presently remain more challenging for other MS protocols. Our investigations lead to the identification of 9 extracellular proteins (e.g. CXCL8, TGFA) and 15 cellular proteins (e.g.TP53, CASP2) in plasma, that were not listed in the plasma PeptideAtlas at the time of our study **Supplementary Excel Sheet: "Antibodies experim. annotation".** For almost 50 % of the polyclonal antibodies annotated as ON- or CO-target in our studies, the identified peptides aligned with the sequence of the protein fragments used to generate the antibodies or Protein Epitope Signature Tags (PrEST) (**Supplementary Fig.1B**). As previously discussed [26], affinity purification of polyclonal antibodies may bare the risk of co-eluting the target used as bait on the columns and detecting these by MS. Such passenger proteins or peptides may consequently simulate the enrichment of an endogenous target. In our case, this would lead to a false classification of the antibody and may hamper downstream applications. To address this concern, we analyzed beads with immobilized antibodies for the presence of the protein fragment antigen that could cause passenger peptides to appear. Out of 42 tested antibodies (among ON-target or CO-target), only 11 showed the potential presence of passenger proteins (**Supplementary note 4, Supplementary Fig. 6)**.

CO-target enrichment and sub-categories

Evaluating antibodies in terms of target selectivity, we observed the possibility to study co-enrichments of proteins (**Figure 3)**. This category includes those ON-target enrichments where other proteins were either enriched due to an interaction with the ON-target or as an OFF-target being captured by the antibodies alongside the indented target. For the second sub-category, sequence homology and abundance of the OFF-target could serve as reasons for co-enrichments.

CO-target enrichment stemming from related proteins

Besides IGFBP2 discussed further below, examples of the CO-target enrichments with sequence homology included examples for anti-CCL16 (HPA042909) and anti-SERPINA4 (HPA002869). These binders respectively enriched additional members of their protein family, CCL18 and SERPINA6, both sharing sequence homology with the intended target (**Supplementary note 7**). The proteins CCL16 and CCL18 are estimated to be present at 29 ng/ml and 2 ng/ml levels in human plasma (Peptide Atlas, 2017), hence the estimated abundance differs from the degrees to which HPA042909 captured CCL16 (z = 10.7) and CCL18 (z = 10.1). Both proteins were otherwise rarely observed in other pIPs (f < 4%), have not been predicted to interact directly (**Supplementary Fig. 7A**) but share sequence to a 27% homology (**Supplementary Fig. 7B**). In the case of HPA002869, the antibody enriched SERPINA4 (z = 8.2) and SERPINA6 (z = 8.1) in heated plasma. Serpins are a large family of blood proteins highly homolog proteins and no direct integration has been predicted (**Supplementary Fig. 7C).** SERPINA4 and A6 are estimated to be present at 17 and 41 µg/ml levels

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and share a 40% sequence similarity (**Supplementary Fig. 7D**). Both were observed in 91% (SERPINA4) and 59% (SERPINA6) of all conducted pIPs with heat-treated plasma (less frequent in untreated plasma). A research performed on PIPs, a database of predicted human protein-protein interactions (http://www.compbio.dundee.ac.uk/www-pips/index.jsp, [27]), revealed a predicted interaction between SERPINA4 and SERPINA6 (total interaction score = 128.0). Further antibodies with independent epitopes are needed to confirm or to exclude the possibility of interactions between these proteins, as we did not find literature supporting a physical interaction between them.

CO-target enrichment stemming from frequent proteins

In another sub-category, we observed that potential protein-protein interactions may also be described from more frequently observed proteins that tend to present with z-scores < 3. Below, we discuss two examples found for IGM, a protein known to be abundant in plasma. The first example is given by CD5 antigen-like (CD5L, HPA026432) and its known interactor IGM. Here, CD5L was detected alongside several immunoglobulins (IGH, IGJ, IGK, IGL, IGM). It is known that CD5L takes part in inflammatory responses during infections or during the process of atherosclerosis, hence binds to the Fc region of IgM through its SRCR domains [30]. Further, the immunoglobulin J chain (IGJ) is known to be required to stabilize the binding of CD5L to IGM, but a direct interaction has not been experimentally observed [31]. Our data supports the idea that association between CD5L, IGM and IGJ can occur: Besides CD5L ([c] = 5.9 µg/ml; z = 5.1), immunoglobulin light chain lambda (IGHV3-23; z = 3.4), kappa (IGKC; z = 3.4), IGM (z = 2.7) and IGJ (z = 2.8). A lower z-score of the co-targets may also possibly indicate that the antibody is more selective for CD5L and less for the additionally identified proteins. Considering the abundance of IGM at around 1 mg/ml and that IGM frequently occurred as contaminant (f = 98%), an increased z-score of IGM in this particular IP may point to a specific enrichment. The second example is given for a monoclonal antibody raised against fibulin 1 (FBLN1) [32]. This antibody was used in both sample types and enriched 22 proteins in heated and 12 in untreated plasma. Even though FBLN1 is a frequent protein (f = 65%) and abundant protein ([c] = 34 µg/ml), it was only enriched in heated plasma (z = 3.7), while other targets such as CAPZA2, IL36G and PLK4I were detected in both preparations (**Supplementary Excel Sheet: "Antibodies against same protein"**). Interestingly, CD5L and IGHM were again found in the same assays, with comparable enrichment values in heated (z = 3.4 and 3.3) compared to untreated plasma (both z = 2.9). This could serve as another indication that CD5L and IGHM are present in a complex in plasma. Further analyses are though needed to determine the mode of co-enrichment, meaning, if the co-target was enriched due to directly interacting with the on-target or due to be serving as an off-target.

Studying protein interaction with paired antibodies

The biologically most interesting sub-category within the CO-target category refers to proteins that co-identified because they presumably interact with the intended target. We chose a limit in z ≥ 3 and LFQ intensity ≥ 10^7 to call potential protein interactions. In general, we observed that most of the

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

consistent identifications (identified in several replicates) were found for LFQ precursor intensities above this level (**Supplementary Fig. 10**). One such example is given by the pIP using three antibodies raised against the insulin growth factor binding protein 2 (IGFBP2: HPA077723, HPA045140, HPA004754) of which the latter two were raised against the same antigen. As shown in **Fig. 3B** for untreated plasma, HPA077723 and HPA045140 both enriched IGFBP2 ([c] = 1.1 µg/ml; *f*=21%) as well as previously known interactors insulin growth factor 1 (IGF1: [c] = 0.46 µg/ml; *f* = 18%;) and IGF2 ([c] =1.6 µg/ml; 8%). In addition, the plasma proteins butyrylcholine esterase (BCHE: [c] = 11.0 µg/ml; *f* = 18%;) and the deoxyribose-phosphate aldolase (DERA: [c] = 0.5 ng/ml; *f* = 7%) were detected. For BCHE and IGF1, an interaction was indeed previously hypothesized [28, 29]. For the third binder (HPA004754), IFGBP2 was only enriched upon prior heat treatment plasma (**Fig. 3B**). This differential performance of antibodies raised against the same antigen (HPA045140, HPA004754) confirms the necessity to investigate each of the different batches and lots of polyclonal antibodies.

In order to provide further evidence for an interaction between the identified proteins, we conducted multiplexed sandwich assays. Here, recombinant IGFBP2, IFG1, IGF2 and BCHE were first analyzed in a concentration dependent manner to confirm assay functionality and target specificity **(Table 2)**. First, EDTA plasma was analyzed in a concentration dependent manner, confirming the selectivity of the matched antibody pairs **(Supplementary Fig. 11)**. Then, we investigated if antibody pairs with different selectivity revealed plasma concentration dependent results. As shown in **Fig 4**, we found pairs of antibodies mixed specificity in the following capture-detection configurations: IGFBP2-IGF2, IGF2-IGFBP2 as well as BCHE-IGFBP2. For IGF1 and IGF2 antibody pairs, it was not possible to obtain a dilution curve with the respective recombinant proteins in solution, but they were functional in plasma (**Supplementary Fig. 11 C,D,M,N,and P)**). Also, IGF2-IGFBP2 and IGFBP2-IGF1 confirmed the presence of the previously known complex IGFBP2-IGF2 (**Table 2**, **Fig 4 B-C**). Antibody pairs for IGFBP2 and BCHE described a sample dilution depended trend with their corresponding intended recombinant proteins as well as in plasma (**Table 2** and **Supplementary Fig. 11 A, E,F,H and I)**). Since no cross-reactivity was observed towards these two proteins with other antibodies in the assay (**Table 2**), the functional antibody pair BCHE-IGFBP2 supports the indications provided by IP that a physical interaction between the two proteins in plasma (**Fig 4A**). An inverted configuration IGFBP2-BCHEand an assay including also the other IGFBP2 antibody indicating an interaction (HPA077723, **Fig. 3B**) would further support this observation. However, we acknowledge that not all antibodies allowed building mixed sandwich pairs and using the chosen assay protocol, and indeed, HPA004754 and HPA077723 were raised against two different regions of IGFBP2. This could explain their different performance as capture and detection, above all in the presence of complexes containing IGFBP2. HPA004754 was though functional as capture and detection antibody both using heat-treated and untreated plasma for the detection of IGFBP2 and in combination with anti-BCHE

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(**Supplementary Fig. 11. B and G**). HPA077723 was though not functional with anti-BCHE either as a capture or detection antibody, suggesting that the binding of either antibody hinders the other antibody binding to IGFBP2-BCHE complex. Further investigations are needed to investigate if this hindrance is due to a proximity of the two binding sites or other steric effects such as epitope accessibility of a captured complex.

More examples of potential protein interactions are provided by those cases in which multiple antibodies were raised against the same target protein using different antigens. Enriching common ON-target and CO-targets provides evidence for protein complexes rather than artefacts. Examples are antibodies targeting FBLN1, or IGF1R (see **Supplementary Excel Sheet: "Antibodies against same protein"**).

OFF-target enrichment category

As a last category, we investigate off-target enrichments. Here, we see the plasma abundance of the off-target over the intended analyte as a main reason for failing to enrich the expected target. As the community is starting to acknowledge the fact that the performance of antibodies is indeed sample context and application dependent, certifying off-target interactions may still allow generating novel hypotheses given that these are followed-up and thoroughly validated by appropriate targeted analysis.

One example for selective off-target interactions in plasma is presented by the antibody HPA004920, raised against MMP1 (Matrix metalloproteinase 1). We classified this antibody as OFF-target in untreated plasma because it enriched Mannose-binding protein C (MBL2; z = 8.3; *f*=12%) as well as MMP3 (z = 6.8; *f*=6%) in the IP assays (Figure 2C). As described above for CCL16 and SERPINA4, also here a 53% sequences similarity between the intended target (MMP1) and the off-target (MMP3) (**Supplementary Fig. 7F**) exists and an interaction between these two proteins has been predicted **(Supplementary Fig. 7E)**. The other off-target MBL2 and MMP1 have only a 10% sequence similarity (**Supplementary Fig. 7G**). MBL2 is though estimated to be present at 1.7 µg/ml in circulation, it is almost 1000x more abundant than MMP1 ([c] = 1.1 ng/ml) and MMP3 ([c] = 0.5 ng/ml) [14]. MBL2 has also been described in to reside in a complex with the serine protease MASP (MBL-associated serine protease) [31], and its collagen-like domain may serve as a substrate for matrix metalloproteinases (MMPs) to nest in. Studies on MBL mutations suggest that MMPs may be involved in physiological regulation of MBL levels [32]. This could eventually explain the presence of MBL2-MMP3 complexes in plasma.

Conclusion

In summary, this study describes a resource that was built from our interest in the verification of antibodies selectivity in plasma. The antibodies analyzed in this study include polyclonal and monoclonal that was used for exploratory bead arrays and the development of immunoassays, where either heat-treated or untreated plasma may serve as samples. We have conducted > 400 IP assays in plasma and built a library of proteins with their frequencies of identification in plasma. Constructed on

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the systematic analysis of 157 antibodies, we described the occurrence of common proteins, denoted plasma background, which allowed us to determine the selectively captured endogenous plasma proteins by mean of z-scores analysis. Our approach, which we also compared with Western blot (**Supplementary note 5** and **Supplementary Table 1**), could serve as a valuable method to narrow large numbers of antibodies determining which ones could enrich the endogenous protein of interest and to further investigate assay selectivity as well as proteins interactions.

This concept may though not yet elucidate (i) if the antibodies bind to proteins at full-length or fragments, (ii) if the antibodies will be functional in pairs in sandwich assay, (iii) how potential protein interactors and off-targets would compete with on-target binding, (iv) how contaminations from passenger antigens affect the assay's selectivity and sensitivity. For ON-target and CO-target antibodies, thus further investigations are required to clarify the technical aspects mentioned above and to expand on the biological implication of protein interactions in plasma. Nevertheless, pIP is an informative first test for the identification of pair antibodies with different target proteins to study protein complexes in plasma by sandwich assays. While, targeted fit-for-purpose experiment should then include dose response curve in dilutions of plasma, preferentially coupled to quantitative mass spectrometry analysis for a set of identified peptides including potential co-targets and/or contaminants.

The provided resource builds one foundation towards a more detailed assessment of antibodies for plasma proteomics assays, and may contribute to the development and application of more specific, robust and reliable immunoassays that can use mass spectrometry or other means of detection [1].

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

MATERIAL AND METHODS

Methods are available in the online version of the paper.

AUTHOR CONTRIBUTIONS

CF and JMS initiated, designed, and led the study with the scientific contribution and support from all co-authors. CF designed the IP experiments and performed the data analysis. CF, SB, LSR and MI executed the IP experiments with support from DT and RMB.LSR conducted the immunoassays. JL provided MS infrastructure. The study was supervised by PN, JL and JMS. CF and JMS wrote the manuscript with input from all authors.

AKNOWLEDGMENTS

We greatly thank Mathias Uhlén, Fredrik Edfors, Björn Forström, Lucia Lourido, Gabriella Tekin and the Biobank profiling, Affinity Proteomics and Clinically Applied Proteomics groups at SciLifeLab in Stockholm for their continuous fruitful discussion, access to instrumentation and input to the presented work. We also thank everyone at the Human Protein Atlas for their support, and thank Hanna Tegel, Johan Rockberg and their team for recombinant proteins. The KTH Center for Applied Precision Medicine (KCAP) funded by the Erling-Persson Family Foundation is acknowledged for financial support. This work was supported by grants for Science for Life Laboratory, the Knut and Alice Wallenberg Foundation. The work leading to this publication has received support from the Innovative Medicines Initiative Joint under grant agreement n°115317 (DIRECT), resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in-kind contribution. Funding from the Swedish Foundation for Strategic Research, Swedish Cancer Society and Swedish Research Council is gratefully acknowledged.

COMPETING FINANCIAL INTERESTS

The authors have on conflicts of interest.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ABBREVIATIONS:

HPA	Human Protein Atlas
LFQ	label free quantification
PrEST	Protein Epitope Signature Tags
Z	z-score
f	frequency of occurrence

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

REFERENCES

1. Baker M. Reproducibility crisis: Blame it on the antibodies. Nature. 2015;521(7552):274-6. doi: 10.1038/521274a. PubMed PMID: 25993940.

2. Uhlen M, Bandrowski A, Carr S, Edwards A, Ellenberg J, Lundberg E, et al. A proposal for validation of antibodies. Nat Methods. 2016;13(10):823-7. doi: 10.1038/nmeth.3995. PubMed PMID: 27595404.

3. Folkersen L, Fauman E, Sabater-Lleal M, Strawbridge RJ, Franberg M, Sennblad B, et al. Mapping of 79 loci for 83 plasma protein biomarkers in cardiovascular disease. PLoS Genet. 2017;13(4):e1006706. doi: 10.1371/journal.pgen.1006706. PubMed PMID: 28369058; PubMed Central PMCID: PMCPMC5393901.

4. Suhre K, Arnold M, Bhagwat AM, Cotton RJ, Engelke R, Raffler J, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. Nature communications. 2017;8:14357. doi: 10.1038/ncomms14357. PubMed PMID: 28240269; PubMed Central PMCID: PMCPMC5333359.

5. Marcon E, Jain H, Bhattacharya A, Guo H, Phanse S, Pu S, et al. Assessment of a method to characterize antibody selectivity and specificity for use in immunoprecipitation. Nat Methods. 2015. doi: 10.1038/nmeth.3472. PubMed PMID: 26121405.

6. Korbakis D, Brinc D, Schiza C, Soosaipillai A, Jarvi K, Drabovich AP, et al. Immunocapture-Selected Reaction Monitoring Screening Facilitates the Development of ELISA for the Measurement of Native TEX101 in Biological Fluids. Mol Cell Proteomics. 2015;14(6):1517-26. doi: 10.1074/mcp.M114.047571. PubMed PMID: 25813379; PubMed Central PMCID: PMC4458717.

7. Korbakis D, Prassas I, Brinc D, Batruch I, Krastins B, Lopez MF, et al. Delineating monoclonal antibody specificity by mass spectrometry. J Proteomics. 2015;114:115-24. doi: 10.1016/j.jprot.2014.11.004. PubMed PMID: 25462431.

8. Anderson NL, Anderson NG, Haines LR, Hardie DB, Olafson RW, Pearson TW. Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). J Proteome Res. 2004;3(2):235-44. PubMed PMID: 15113099.

9. Anderson NL, Jackson A, Smith D, Hardie D, Borchers C, Pearson TW. SISCAPA peptide enrichment on magnetic beads using an in-line bead trap device. Mol Cell Proteomics. 2009;8(5):995-1005. doi: 10.1074/mcp.M800446-MCP200. PubMed PMID: 19196707; PubMed Central PMCID: PMC2689780.

10. Ippoliti PJ, Kuhn E, Mani DR, Fagbami L, Keshishian H, Burgess MW, et al. Automated Microchromatography Enables Multiplexing of Immunoaffinity Enrichment of Peptides to Greater than 150 for Targeted MS-Based Assays. Anal Chem. 2016;88(15):7548-55. doi: 10.1021/acs.analchem.6b00946. PubMed PMID: 27321643.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

11. Li H, Popp R, Frohlich B, Chen MX, Borchers CH. Peptide and Protein Quantification Using Automated Immuno-MALDI (iMALDI). Journal of visualized experiments : JoVE. 2017;(126). doi: 10.3791/55933. PubMed PMID: 28872133.

12. Krastins B, Prakash A, Sarracino DA, Nedelkov D, Niederkofler EE, Kiernan UA, et al. Rapid development of sensitive, high-throughput, quantitative and highly selective mass spectrometric targeted immunoassays for clinically important proteins in human plasma and serum. Clin Biochem. 2013;46(6):399-410. doi: 10.1016/j.clinbiochem.2012.12.019. PubMed PMID: 23313081; PubMed Central PMCID: PMCPMC3779129.

13. Schwenk JM, Omenn GS, Sun Z, Campbell DS, Baker MS, Overall CM, et al. The Human Plasma Proteome Draft of 2017: Building on the Human Plasma PeptideAtlas from Mass Spectrometry and Complementary Assays. J Proteome Res. 2017. doi: 10.1021/acs.jproteome.7b00467. PubMed PMID: 28938075.

14. Fredolini C, Bystrom S, Pin E, Edfors F, Tamburro D, Iglesias MJ, et al. Immunocapture strategies in translational proteomics. Expert Rev Proteomics. 2016;13(1):83-98. doi: 10.1586/14789450.2016.1111141. PubMed PMID: 26558424; PubMed Central PMCID: PMCPMC4732419.

15. Ayoglu B, Haggmark A, Neiman M, Igel U, Uhlen M, Schwenk JM, et al. Systematic antibody and antigen-based proteomic profiling with microarrays. Expert Rev Mol Diagn. 2011;11(2):219-34. doi: 10.1586/erm.10.110. PubMed PMID: 21405972.

16. Neiman M, Fredolini C, Johansson H, Lehtio J, Nygren PA, Uhlen M, et al. Selectivity analysis of single binder assays used in plasma protein profiling. Proteomics. 2013;13(23-24):3406-10. doi: 10.1002/pmic.201300030. PubMed PMID: 24151238.

17. ten Have S, Boulon S, Ahmad Y, Lamond AI. Mass spectrometry-based immuno-precipitation proteomics - the user's guide. Proteomics. 2011;11(6):1153-9. doi: 10.1002/pmic.201000548. PubMed PMID: 21365760; PubMed Central PMCID: PMCPMC3708439.

18. Mellacheruvu D, Wright Z, Couzens AL, Lambert JP, St-Denis NA, Li T, et al. The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nat Methods. 2013;10(8):730-6. doi: 10.1038/nmeth.2557. PubMed PMID: 23921808; PubMed Central PMCID: PMCPMC3773500.

19. Keilhauer EC, Hein MY, Mann M. Accurate Protein Complex Retrieval by Affinity Enrichment Mass Spectrometry (AE-MS) Rather than Affinity Purification Mass Spectrometry (AP-MS). Mol Cell Proteomics. 2015;14(1):120-35. doi: 10.1074/mcp.M114.041012. PubMed PMID: 25363814; PubMed Central PMCID: PMC4288248.

20. Schwenk JM, Igel U, Neiman M, Langen H, Becker C, Bjartell A, et al. Toward next generation plasma profiling via heat-induced epitope retrieval and array-based assays. Mol Cell Proteomics. 2010;9(11):2497-507. Epub 2010/08/05. doi: 10.1074/mcp.M110.001560. PubMed PMID: 20682762; PubMed Central PMCID: PMC2984230.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

21. Chen Y, Mao H, Zhang X, Gong Y, Zhao N. Thermal conformational changes of bovine fibrinogen by differential scanning calorimetry and circular dichroism. Int J Biol Macromol. 1999;26(2-3):129-34. PubMed PMID: 10517519.

22. Marx G, Mou X, Hotovely-Salomon A, Levdansky L, Gaberman E, Belenky D, et al. Heat denaturation of fibrinogen to develop a biomedical matrix. J Biomed Mater Res B Appl Biomater. 2008;84(1):49-57. doi: 10.1002/jbm.b.30842. PubMed PMID: 17471522.

23. Boehm TK, Sojar H, Denardin E. Concentration-dependent effect of fibrinogen on IgG-specific antigen binding and phagocytosis. Cell Immunol. 2010;263(1):41-8. Epub 2010/02/24. doi: 10.1016/j.cellimm.2010.02.014. PubMed PMID: 20303075; PubMed Central PMCID: PMCPMC2862818.

24. Boehm TK, DeNardin E. Fibrinogen binds IgG antibody and enhances IgG-mediated phagocytosis. Hum Antibodies. 2008;17(3-4):45-56. PubMed PMID: 19029661.

25. Katafuchi T, Esterházy D, Lemoff A, Ding X, Sondhi V, Kliewer SA, et al. Detection of FGF15 in plasma by stable isotope standards and capture by anti-peptide antibodies and targeted mass spectrometry. Cell Metab. 2015;21(6):898-904. doi: 10.1016/j.cmet.2015.05.004. PubMed PMID: 26039452; PubMed Central PMCID: PMCPMC4454892.

26. Whiteaker JR, Zhao L, Abbatiello SE, Burgess M, Kuhn E, Lin C, et al. Evaluation of large scale quantitative proteomic assay development using peptide affinity-based mass spectrometry. Mol Cell Proteomics. 2011;10(4):M110 005645. doi: 10.1074/mcp.M110.005645. PubMed PMID: 21245105; PubMed Central PMCID: PMCPMC3069346.

27. McDowall MD, Scott MS, Barton GJ. PIPs: human protein-protein interaction prediction database. Nucleic Acids Res. 2009;37(Database issue):D651-6. Epub 2008/11/06. doi: 10.1093/nar/gkn870. PubMed PMID: 18988626; PubMed Central PMCID: PMCPMC2686497.

28. Jara M, Schulz A, Malinowski M, Puhl G, Lock JF, Seehofer D, et al. Growth hormone/insulin-like growth factor 1 dynamics in adult living donor liver transplantation. Liver Transpl. 2014;20(9):1118-26. Epub 2014/08/04. doi: 10.1002/lt.23922. PubMed PMID: 24889799.

29. Durrant AR, Tamayev L, Anglister L. Serum cholinesterases are differentially regulated in normal and dystrophin-deficient mutant mice. Front Mol Neurosci. 2012;5:73. Epub 2012/06/19. doi: 10.3389/fnmol.2012.00073. PubMed PMID: 22723768; PubMed Central PMCID: PMCPMC3378013.

30. Arai S, Maehara N, Iwamura Y, Honda S, Nakashima K, Kai T, et al. Obesity-associated autoantibody production requires AIM to retain the immunoglobulin M immune complex on follicular dendritic cells. Cell Rep. 2013;3(4):1187-98. Epub 2013/04/04. doi: 10.1016/j.celrep.2013.03.006. PubMed PMID: 23562157.

31. Sanjurjo L, Aran G, Roher N, Valledor AF, Sarrias MR. AIM/CD5L: a key protein in the control of immune homeostasis and inflammatory disease. J Leukoc Biol. 2015;98(2):173-84. Epub 2015/06/05. doi: 10.1189/jlb.3RU0215-074R. PubMed PMID: 26048980.

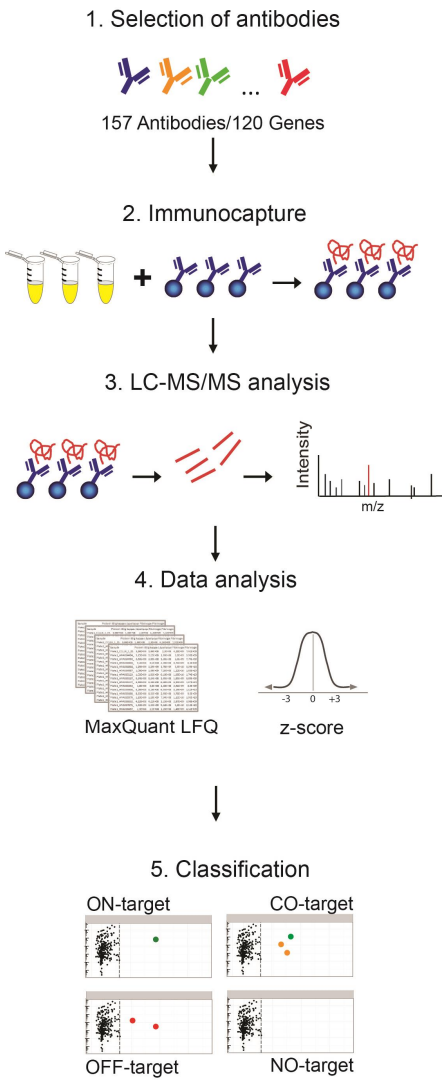
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

32. Neiman M, Hedberg JJ, Donnes PR, Schuppe-Koistinen I, Hanschke S, Schindler R, et al. Plasma profiling reveals human fibulin-1 as candidate marker for renal impairment. J Proteome Res. 2011;10(11):4925-34. doi: 10.1021/pr200286c. PubMed PMID: 21888404.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

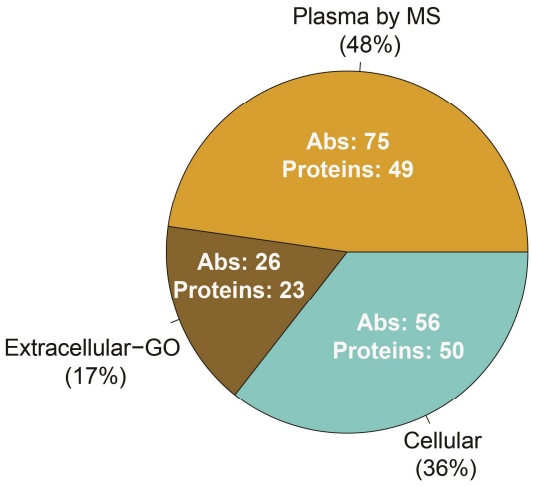
FIGURES & TABLES

Figure 1A: Workflow and study overview: A set of 157 antibodies targeting 120 genes were covalently coupled to magnetic beads and incubated one-by-one with EDTA plasma. Two to four replicate incubations were performed for each antibody. Following target enrichment, washing and digestion on beads, the obtained data files from LC-MS were searched and normalized by MaxLFQ, z-score analysis was performed to rank proteins specifically enriched by each antibody. Using the resource generated by > 400 IP assays, antibodies were classified based on their enrichment profiles: (1) ON-TARGET, only the target protein was enriched showing a z-score ÓB; (2) CO-TARGET, the target protein was enriched together with other proteins also associated to a z-score ÓB; (3) OFF-TARGET, only proteins other than the expected target were enriched; as well as (4) NO-TARGET, in case no protein was enriched (z-scores < 3).



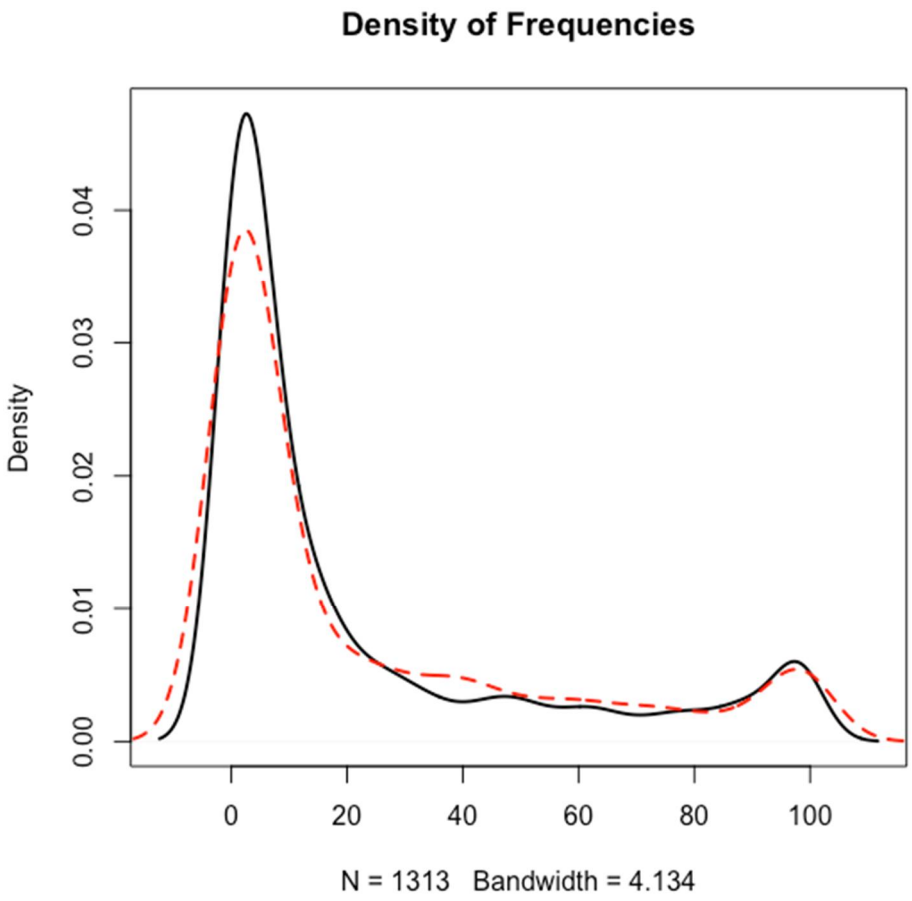
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1B: Distribution of antigen annotation. The target proteins of the 157 antibodies were grouped as: ñPlasma by MSò, were identified in plasma previously by mass spectrometry as reported by Peptide Atlas. Cellular and Extracellular were assigned according to Gene Ontology classification (see Materials and Methods). Numbers stated inside the pie chart refer to the number of antibodies (Abs) in the category and corresponding number of target proteins..



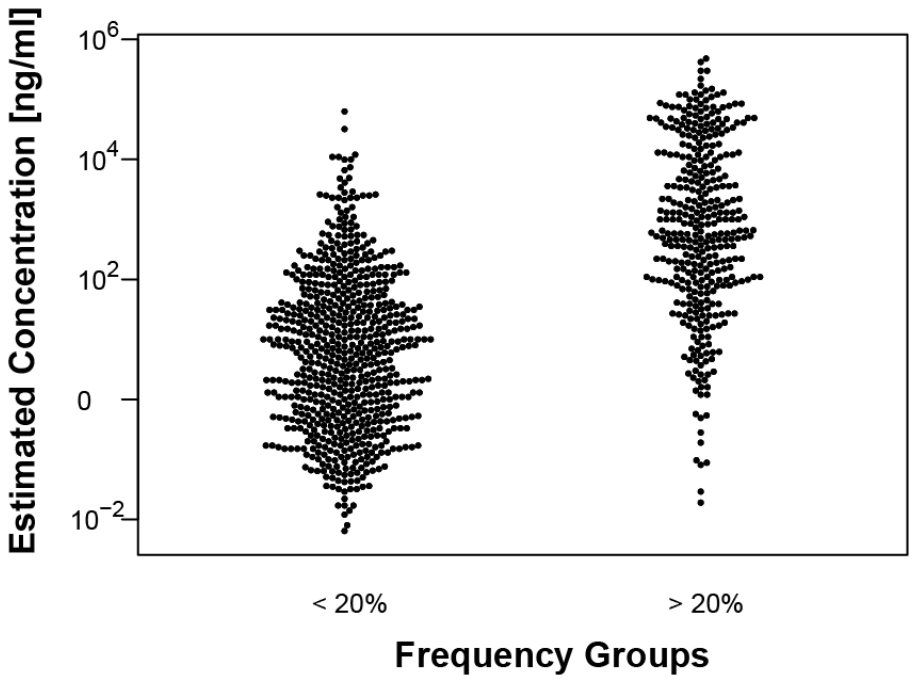
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1C: Distribution of frequencies of identification. The proteins obtained from assays conducted in heat-treated (red) vs untreated plasma (black) were collected in terms of the number of times they were observed in the IP-MS data. For both sample types, the majority of the 1313 proteins were found in less than 20% of the IPs.



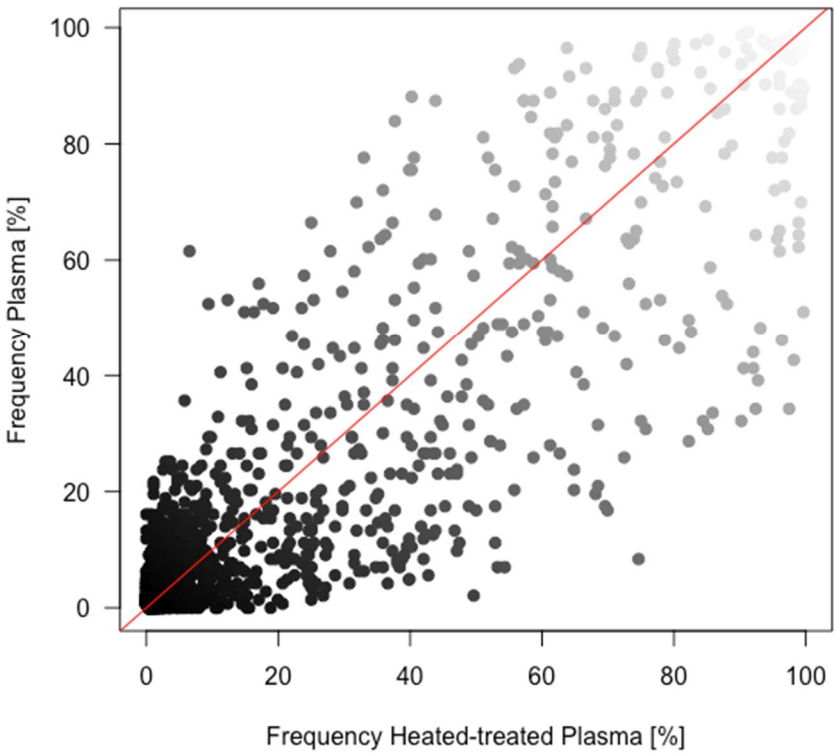
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1D: Frequency vs Concentration. Estimated concentrations in PeptideAtlas were compared between frequent (>20%) and less frequent (<20) protein identifications.



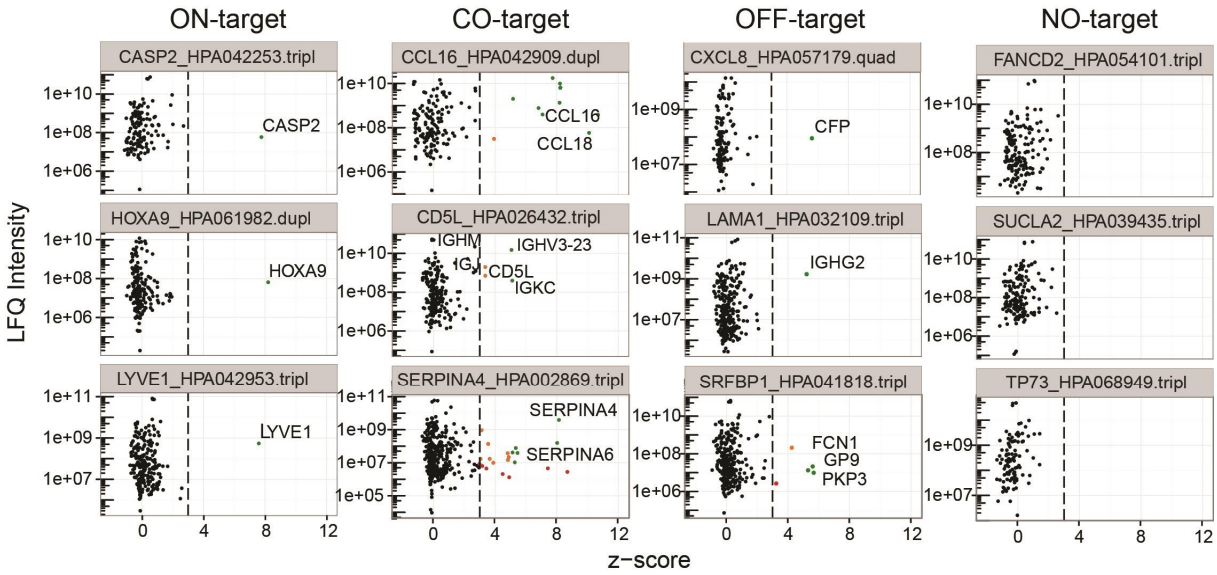
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 2: Frequency correlation. The relation between the frequency percentages of protein occurrence in IPs performed with heat-treated and untreated plasma is shown. The red line represents the line of identity.



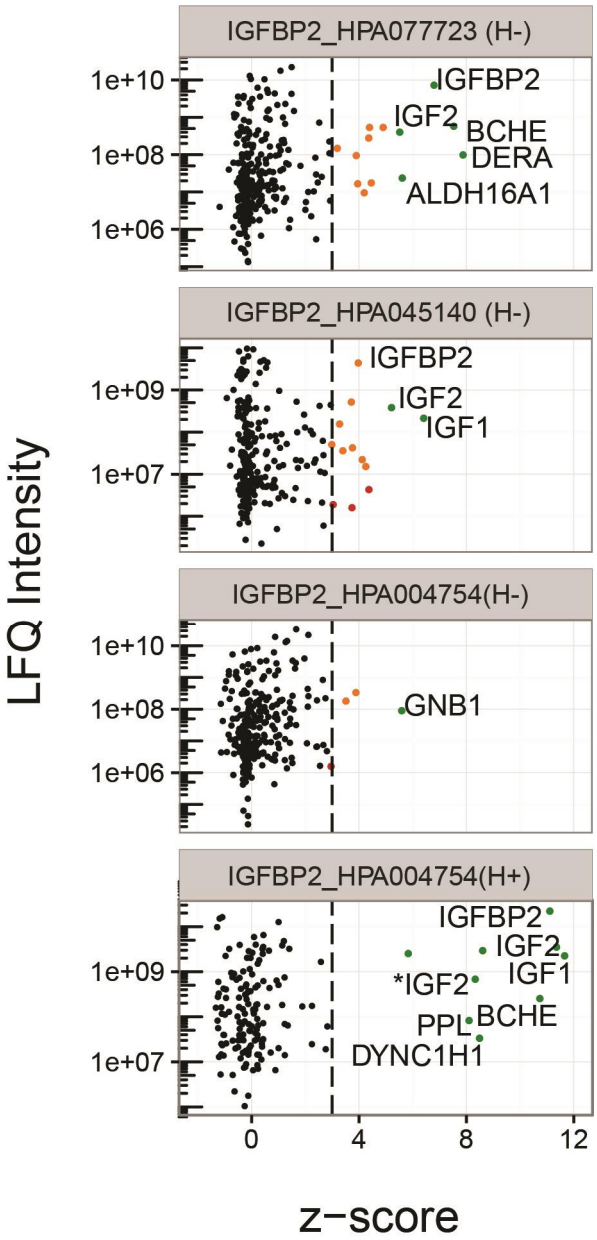
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 3 A: Classification of antibodies. Three representative examples of are given for each of the enrichment categories (ON-target, CO-target, OFF-target, and NO-target). On the top of each plot are the target gene, antibody ID and the number of replicates IP performed for the antibody. The dots in each plot represent identifications present in all the replicates available for the specific antibody.Green dots: z-score >5 and LFQ intensity > 1e+07; yellow dots: z-score >3 and LFQ intensity > 1e+07; red dots: z-score >3 and LFQ intensity < 1e+07. Text: expected target and hypothesized off-targets or interactors. A complete list protein identified and relative z-scores are available in **Supplementary Excel Table**, Sheet **ñz-score> 2.5ò.**



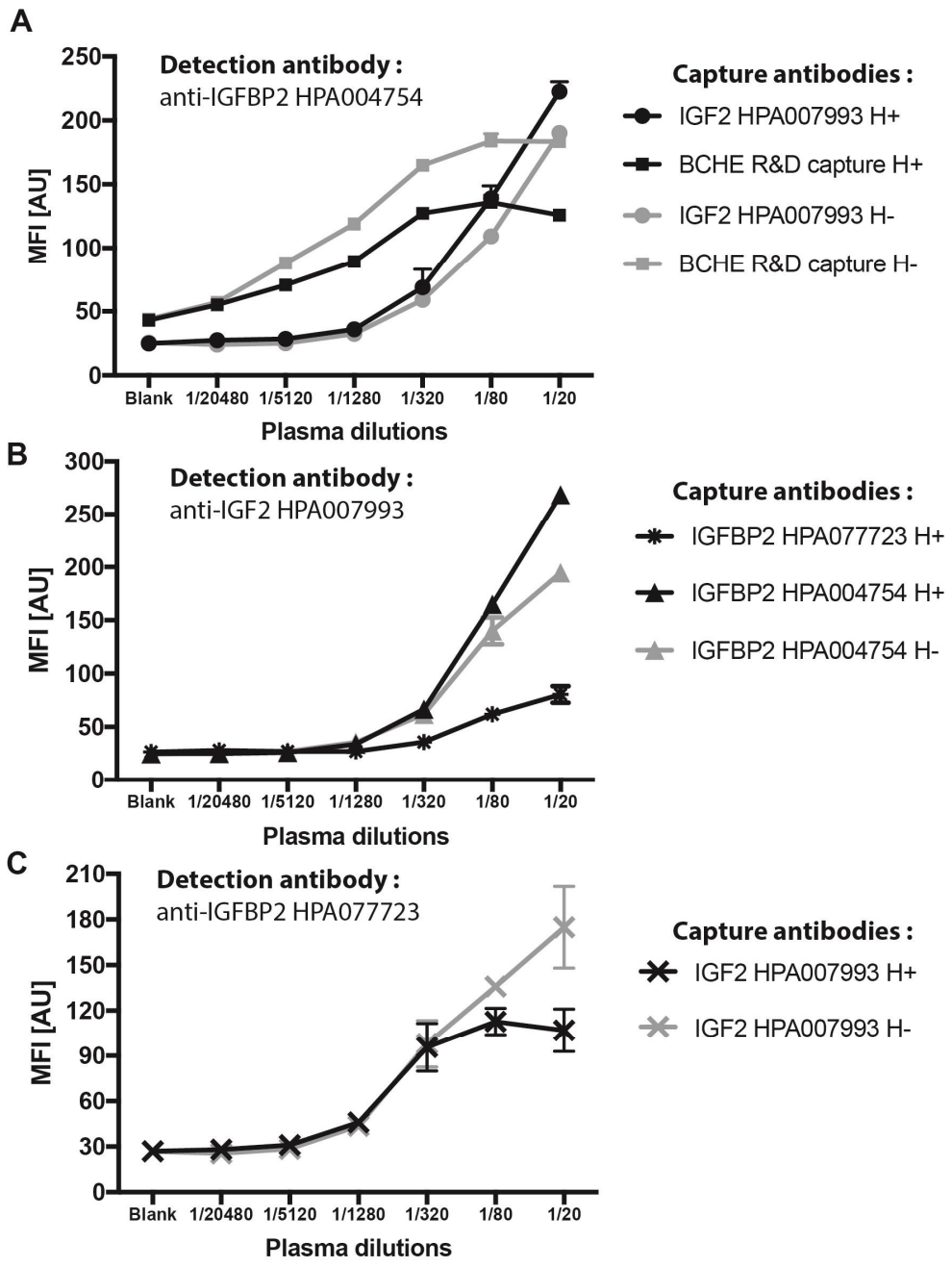
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 3 B: Paired antibodies and co-enrichment profiles. The z-score/LFQ intensity plots of paired antibodies raised against IGFBP2 are shown for HPA004754, HPA045140, HPA077723. (H+) refers to heat treated plasma and (H-)to untreated plasma . IGF2 was identified as P01344, and P01344-2 (*).



bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 4:IGFBP2 protein interaction analysis by sandwich immunoassay. Dilution curves of plasma analyzed by sandwich assays using different combination of capture and detection antibodies. Dots represents mean value with standard error (SD) bars. In black, heated plasma (H+); in gray,non-heated plasma (H-).



bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table 1: Annotation and categorization and of antibodies.

Annotated target location	On/Co-Target		Off/No-Target		Sum	
	N	%	N	%	N	%
Cellular	15	27	41	73	56	36
Extracellular	9	34	17	65	26	16
Plasma (MS)	46	61	29	39	75	48
Total	70	45	87	55	157	100

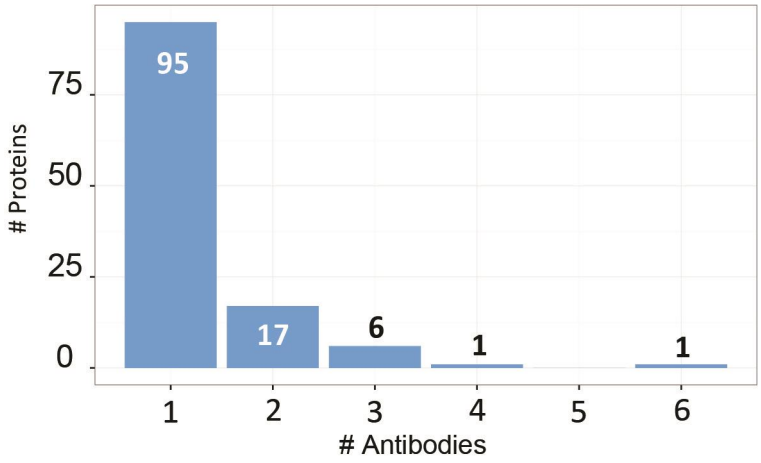
Table 2:Antibody pairs tested in plasma and with recombinant proteins. Annotation (✓): Trends from sample dilution assays were obtained with at least one combination of antibodies for the same protein either in heated or untreated samples. Annotation (-): No sample concentration dependent data was obtained. (Details regarding each single pair in Supplementary Excel Sheet ñReagents_Lot_Numbersò, catalog numbers of the functional pairs are indicated in **Figure 4** and **Supplementary Figure 10**.

SAMPLE	DETECTION ANTIBODY	CAPTURE ANTIBODIES				
		anti-IGFBP2	anti-IGF2	anti-BCHE	anti-IGF1	anti-DERA
Plasma	**anti-IGFBP2**	✓	✓	✓	-	-
	anti-IGF2	✓	✓	-	-	-
	anti-BCHE	-	-	✓	-	-
	anti-IGF1	-	-	-	✓	-
	anti-DERA	-	-	-	-	-
Rec-IGFBP2	**anti-IGFBP2**	✓	-	-	-	-
Rec-IGF2	**anti-IGF2**	-	-	-	-	-
Rec-BCHE	**anti-BCHE**	-	-	✓	-	-
Rec-IGF1	**anti-IGF1**	-	-	-	-	-

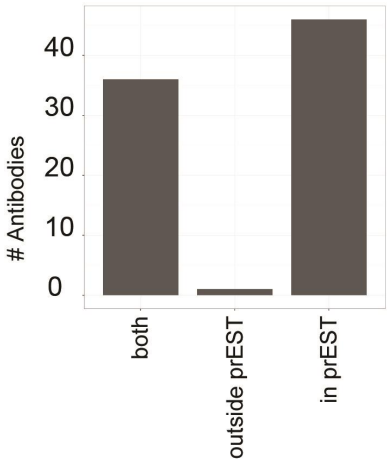
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

SUPPLEMENTARY FIGURES & TABLES

Supplementary Figure 1A: Distribution of number of antibodies targeting each protein.

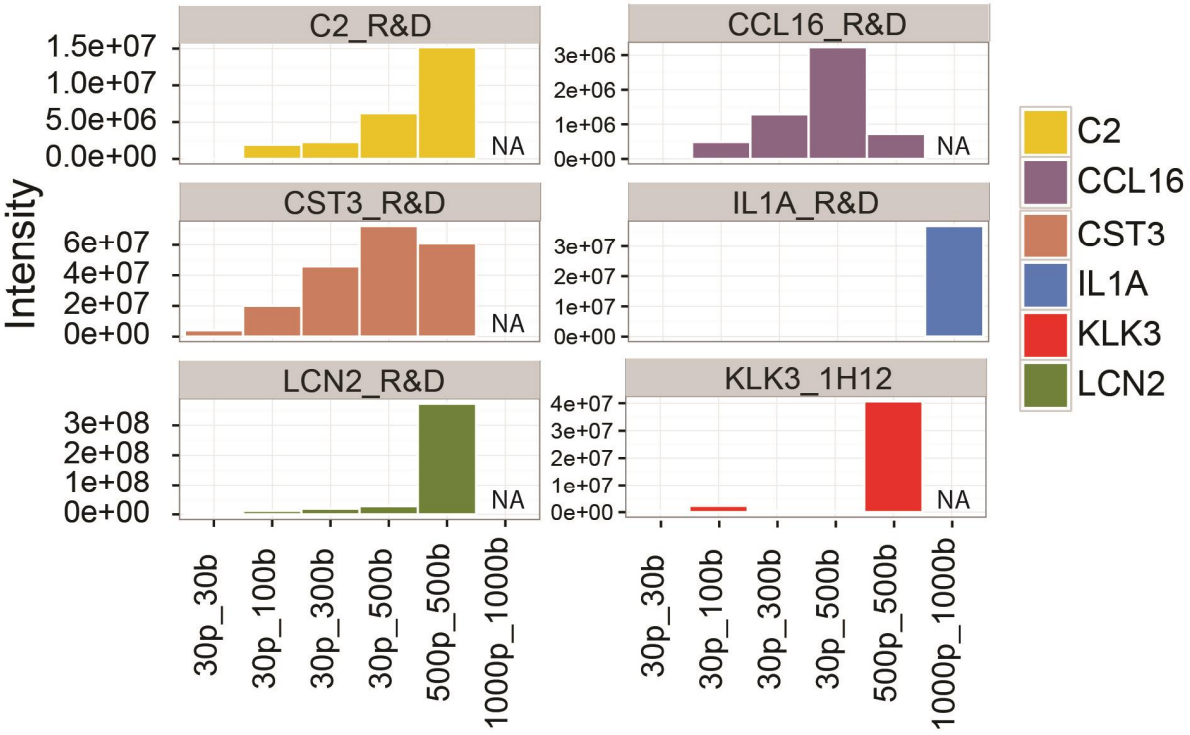


Supplementary Figure 1B: Coverage of proteins identified. Distribution of ON-target and CO-target antibodies based on the peptides identified for the intended target: peptides only covering the PrEST sequence (in PrEST), peptides not covering the PrEST (outside PrEST) or peptides identified inside and outside PrEST sequence (both).



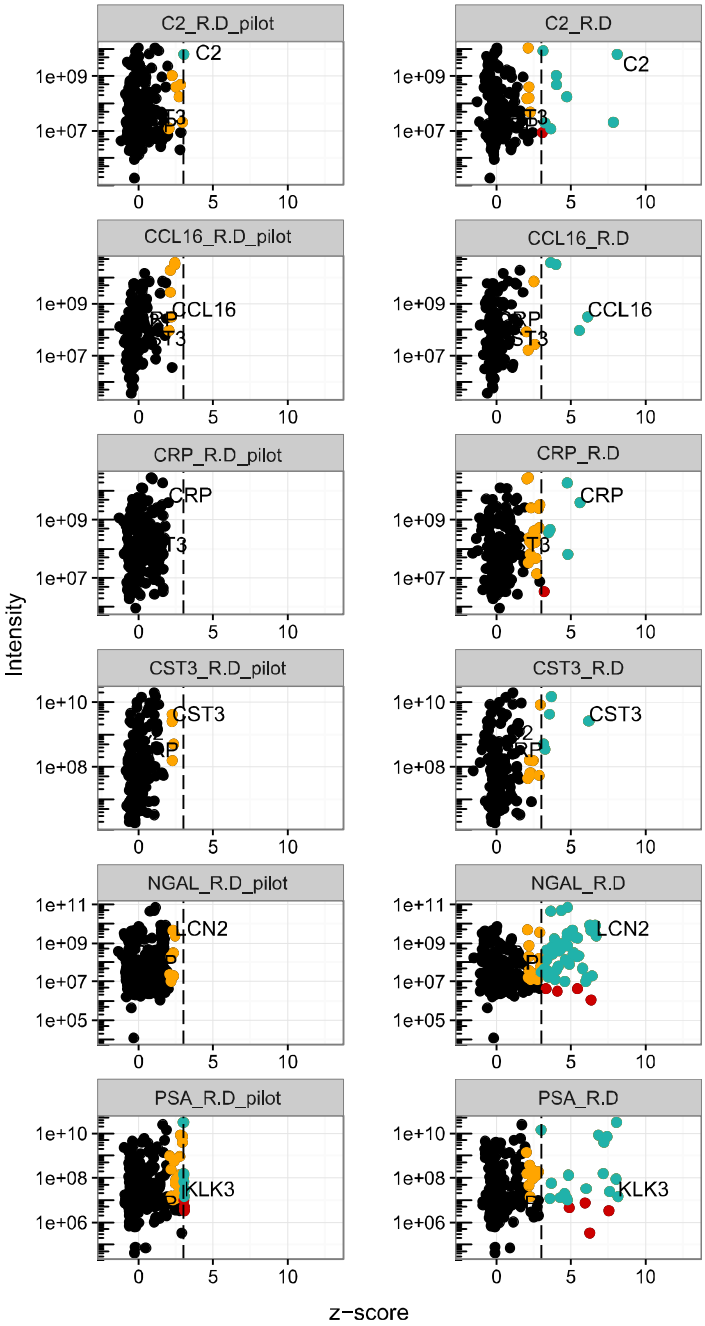
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 2A: Optimization of experimental and technical conditions. Commercial antibodies targeting known plasma proteins of different abundance (C2, CCL16, CST3, IL1A, KLK3, LCN2) were used to establish optimal experimental conditions for volume of plasma and number of beads coupled to antibody. ñpò=volume of plasma in µL; ñbò= number of coupled beads (500000b = 1.6 µg of antibody).



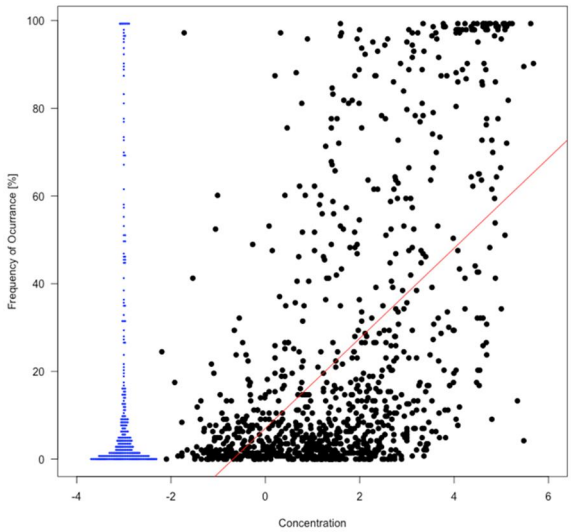
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 2B: Effect of library size. Data derived from IP assays with antibodies targeting the plasma proteins of C2, CCL16, CRP, CST3, NGAL and KLK3 (PSA) was analyzed using the pilot library limited to only these 21 IPs (left column,òpilotò) as well as the library of more than 400 IP (right**).** In this experiment, 100 µl of plasma and 500,000 beads were applied.

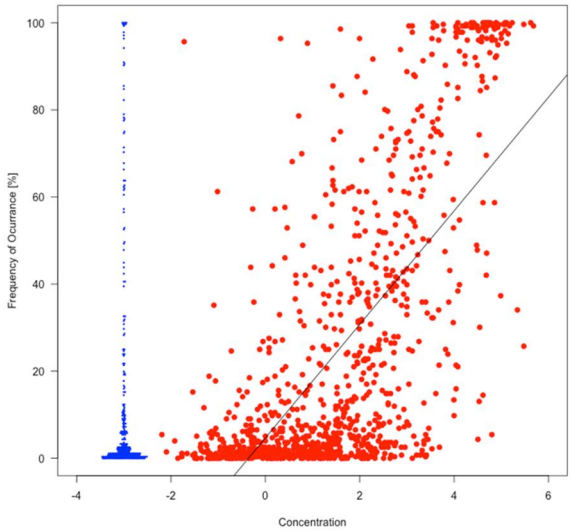


bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 3 A: Relation between estimated plasma concentration [ng/ml] and frequency [%] for untreated plasma. Blue dots represent proteins for which estimated values of concentration were not present in PeptideAtlas.



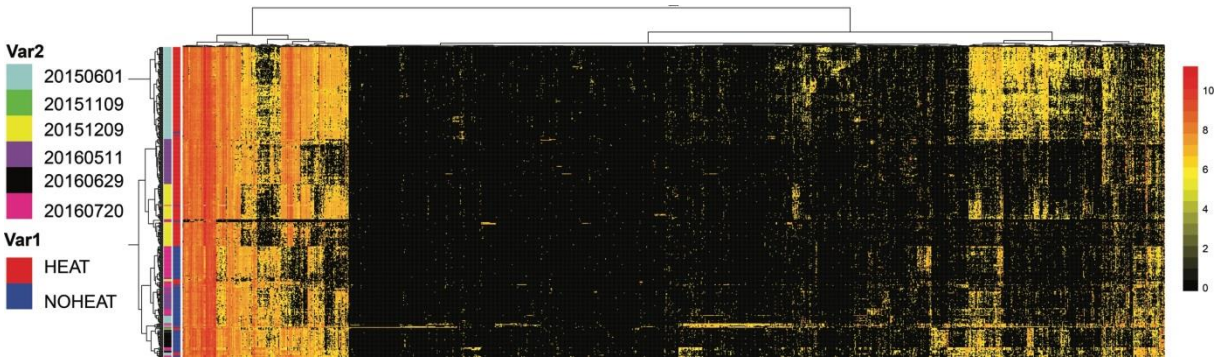
Supplementary Figure 3 B: Relation between estimated plasma concentration [ng/ml] and frequency [%] for heat-treated plasma. Blue dots represent proteins for which estimated values of concentration were not present in PeptideAtlas.



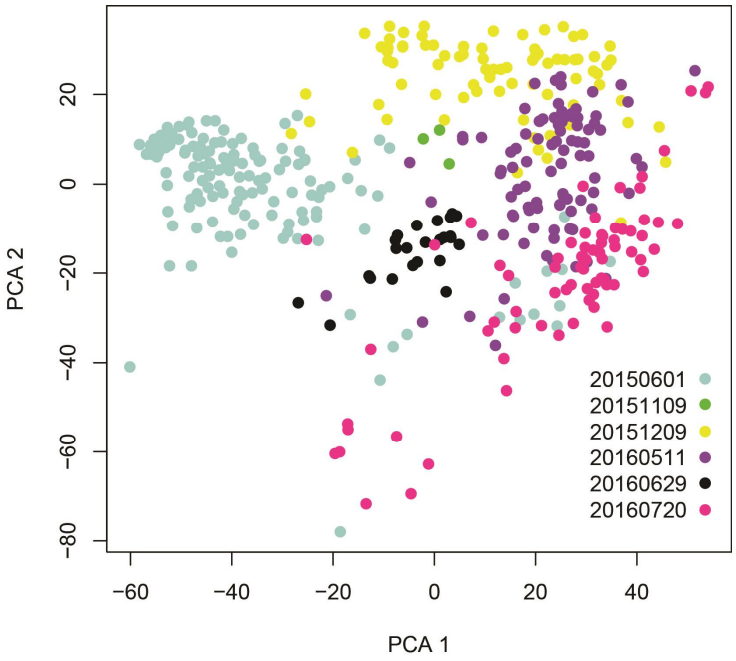
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 4 Evaluation of contaminant proteins identified by pIPs and comparison of experimental batches.

Supplementary Figure 4 A: Two ways hierarchical clustering analysis of the proteins identified in 414 immunoprecipitations. Clustering distance: "euclideanò; Clustering method: Ward. Cluster bars represent (1) different experimental batches and (2) sample heat treatment.

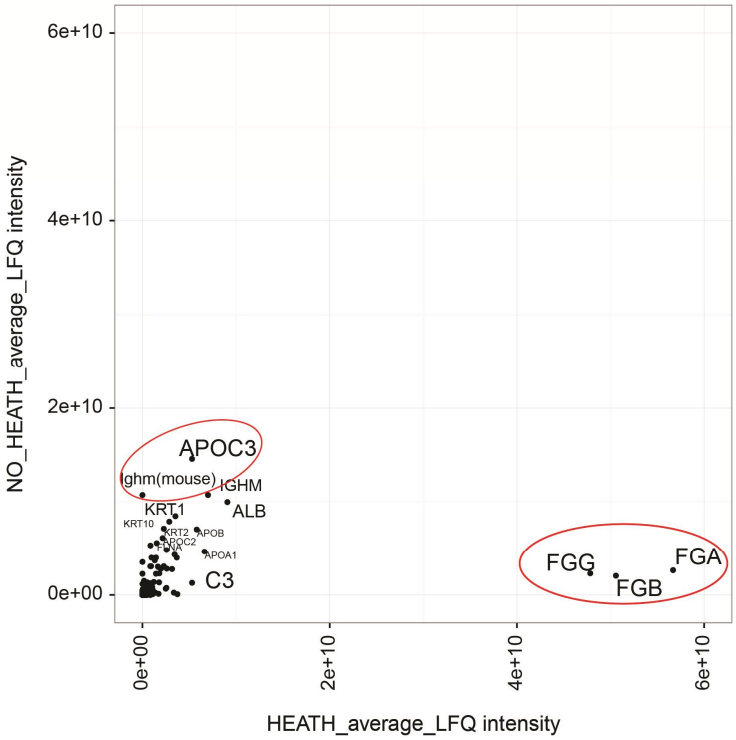


Supplementary Figure 4 B: Representation of principal component analysis (PCA) with dots colors indicating the experimental batches.

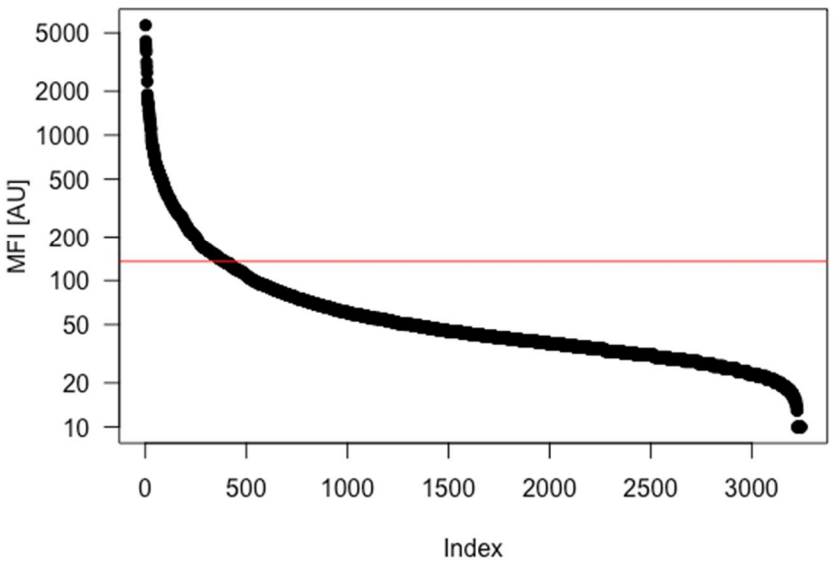


bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 5. Comparison of LFQ protein intensities for IMS performed in heat treated versus in not heat treated plasma. Circled in red high abundant plasma proteins which abundance in the background increases when plasma is heat treated.

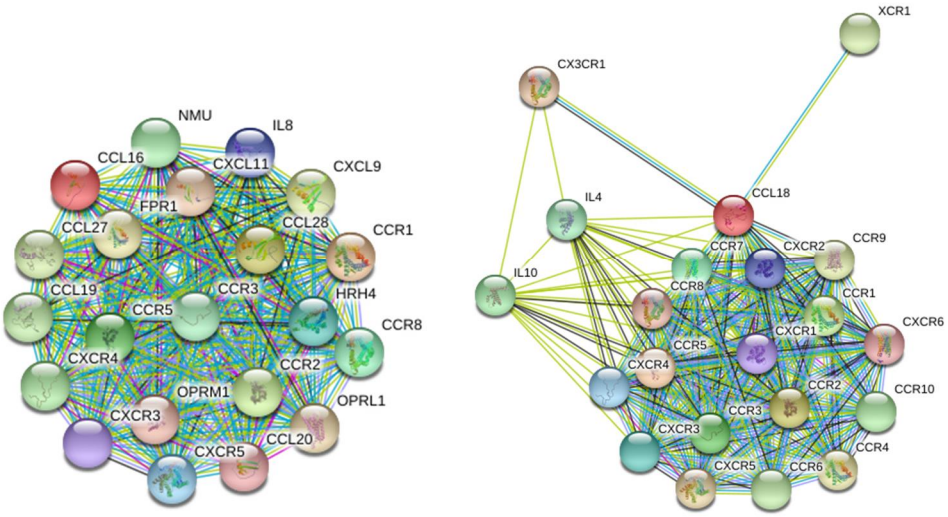


Supplementary Figure 6: Ranked distribution of the detection of passenger antigens on antibody coupled beads**.**



bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 7A: Interaction networks of CCL16 and CCL18 taken from String database V10.5 (20 interactors, 1st shell only).

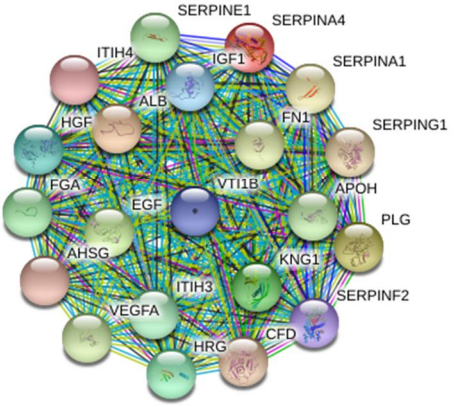


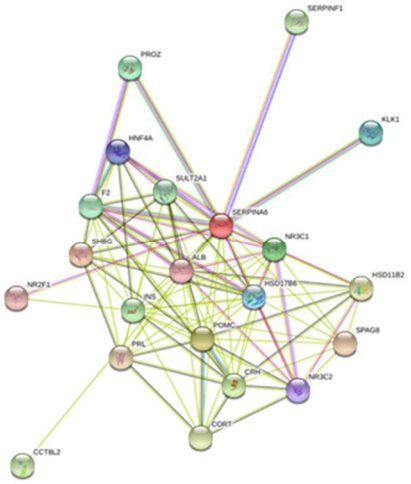
Supplementary Figure 7B: Sequence homology search of CCL16-CCL18 using CLUSTALO alignment revealed a 26.6% homology. In Yellow: Amino acidic sequence used as antigen in the generation of the antibody (PrEST).

```
CLUSTAL O(1.2.3) multiple sequence alignment

SP|O15467|CCL16_HUMAN MKVSEAALSLLVLILIITSASRSQPKVPEWVNTPSTCCLKYYEKVLPRRLVVGYRK-ALN 59
SP|P55774|CCL18_HUMAN MKGLAAA--LLVLVCTMALC-----SCAQVGTNKELCCLVYTSWQIPQKFIVDYSETSPQ 53
                      **   **  ****:  :: .     .  :  .. . *** * .  :*::::*.* : : :

SP|O15467|CCL16_HUMAN CHLPAIIFVTKRNREVCTNPNDDWVQEYIKDPNLPLLPTRNLSTVKIITAKNGQPQLLNS 119
SP|P55774|CCL18_HUMAN CPKPGVILLTKRGRQICADPNKKWVQKYISDLKLNA------------------------ 89
                      *  *.:*::***.*::*::**..***:**.* :*

SP|O15467|CCL16_HUMAN Q 120
SP|P55774|CCL18_HUMAN -
```

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 7C: Interaction networks of SERPINA4 and SERPINA6 taken from String database V10.5 (20 interactors, 1st shell only).



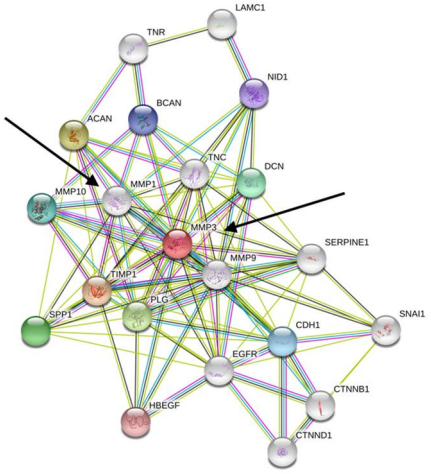


Supplementary Figure 7D: Sequence homology search of SERPINA4-SERPINA6 using CLUSTALO alignment revealed a 40.0% homology. In Yellow: Amino acidic sequence used as antigen in the generation of the antibody (PrEST).

```
CLUSTAL O(1.2.3) multiple sequence alignment

SP|P29622|KAIN_HUMAN MHLIDYLLLLLVGLLALSHGQLHVEHDGESCSNSSHQQILETGEG SPSLKIAPANADFAF 60
SP|P08185|CBG_HUMAN  MPLLLYTCLLWLP----TSGLWTVQAMDP--------NAAYVNMSNHHRGLASANVDFAF 48
                     * *: *  ** :     : *   *:  .         :   .. ..    :* **.****

SP|P29622|KAIN_HUMAN RFYYLIASETPGKNIFFSPLSISAAYAMLSLGACSHSRSQILEGL GFNLTELSESDVHRG 120
SP|P08185|CBG_HUMAN  SLYKHLVALSPKKNIFISPVSISMALAMLSLGTCGHTRAQLLQGL GFNLTERSETEIHQG 108
                      :*  :.: :* ****:**:*** * ******:*.*:*:*:*:*:******* **:::*:*

SP|P29622|KAIN_HUMAN FQHLLHTLNLPGHGLETRVGSALFLSHNLKFLAKFLNDTMAVYE AKLFHTNFYDTVGTIQ 180
SP|P08185|CBG_HUMAN  FQHLHQLFAKSDTSLEMTMGNALFLDGSLELLESFSADIKHYYES EVLAMNFQDWATASR 168
                     **** : :   . .**  :*.****. .*::* .*  *    **::::  ** * . : :

SP|P29622|KAIN_HUMAN LINDHVKKETRGKIVDLVSELKKDVLMVLVNYIYFKALWEKPFI SSRTTPKDFYVDENTT 240
SP|P08185|CBG_HUMAN  QINSYVKNKTQGKIVDLFSGLDSPAILVLVNYIFFKGTWTQPFDL ASTREENFYVDETTV 228
                      **.:**::*:******.* *.. .::******:**. * :**  : *  ::*****.*.

SP|P29622|KAIN_HUMAN VRVPMMLQDQEHHWYLHDRYLPCSVLRMDYKGDATVFFILPNQ GKMREIEEVLTPEMLMR 300
SP|P08185|CBG_HUMAN  VKVPMMLQSST-ISYLHDSELPCQLVQMNYVGNGTVFFILPDKG KMNTVIAALSRDTINR 287
                     *:******..    ****  ***.:::*:* *:.*******::***. :  .*: : : *

SP|P29622|KAIN_HUMAN WNNLLRKRNFYKKLELHLPKFSISGSYVLDQILPRLGFTDLFSKW ADLSGITKQQKLEAS 360
SP|P08185|CBG_HUMAN  WSAGLTS----SQVDLYIPKVTISGVYDLGDVLEEMGIADLFT NQANFSRITQDAQLKSS 343
                     *.  * .    .:::*::**.:*** * *.::* .:*::***:: *::* **:: :*::*

SP|P29622|KAIN_HUMAN KSFHKATLDVDEAGTEAAAATSFAIKFFSAQTNRHILRFNRPFLV VIFSTSTQSVLFLGK 420
SP|P08185|CBG_HUMAN  KVVHKAVLQLNEEGVDTAGSTGVTLNL---TSKPIILRFNQPFIIMIFDHFTWSSLFLAR 400
                     * .***.*:::* *.::*.:*..::::    ::  *****:**:::**.  * * ***.:

SP|P29622|KAIN_HUMAN VVDPTKP 427
SP|P08185|CBG_HUMAN  VMNPV-- 405
                     *::*.
```

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 7E: Interaction network of MMP1 and MMP3 taken from String database V10.5 (20 interactors, 1st shell only).



Supplementary Figure 7F: Sequence homology search of MMP1-MMP3 using CLUSTALO alignment revealed a 52.8 % homology. CLUSTAL. In Yellow: Amino acidic sequence used as antigen in the generation of the antibody(PrEST)..

```
CLUSTAL O(1.2.3) multiple sequence alignment

SP|P03956|MMP1_HUMAN MHSFPPLLLLLFWGVVSHSFPATLETQEQDVDLVQKYLEKYYNLKNDGRQVEKRRNSGPV 60
SP|P08254|MMP3_HUMAN MKSLPILLLLCVAVCSAYPLDGAARGEDTSMNLVQKYLENYYDLKKDVKQFVRRKDSGPV 60
                     *:*:* **** .    :: : .: . :: .::*******:**:**:* :*. :*::****

SP|P03956|MMP1_HUMAN VEKLKQMQEFFGLKVTGKPDAETLKVMKQPRCGVPDVAQFVLTEGNPRWEQTHLTYRIEN 120
SP|P08254|MMP3_HUMAN VKKIREMQKFLGLEVTGKLDSDTLEVMRKPRCGVPDVGHFRTFPGIPKWRKTHLTYRIVN 120
                     *:*:::**:*:**:**** *::**:**::********.:*    * *:*.:******* *

SP|P03956|MMP1_HUMAN YTPDLPRADVDHAIEKAFQLWSNVTPLTFTKVSEGQADIMISFVRGDHRDNSPFDGPGGN 180
SP|P08254|MMP3_HUMAN YTPDLPKDAVDSAVEKALKVWEEVTPLTFSRLYEGEADIMISFAVREHGDFYPFDGPGNV 180
                     ******:  ** *:***:::*.:******::: **:*******.  :* *  ******.

SP|P03956|MMP1_HUMAN LAHAFQPGPGIGGDAHFDEDERWTNNFREYNLHRVAAHELGHSLGLSHSTDIGALMYPSY 240
SP|P08254|MMP3_HUMAN LAHAYAPGPGINGDAHFDDDEQWTKDTTGTNLFLVAAHEIGHSLGLFHSANTEALMYPLY 240
                     ****: *****.******:**:**::    **. *****:****** **::  ***** *

SP|P03956|MMP1_HUMAN TFSGD---VQLAQDDIDGIQAIYGRSQNP---------VQPIGPQTPKACDSKLTFDAIT 288
SP|P08254|MMP3_HUMAN HSLTDLTRFRLSQDDINGIQSLYGPPPDSPETPLVPTEPVPPEPGTPANCDPALSFDAVS 300
                         *   .:*:****:***::**   :            *  * **  **  *:***::

SP|P03956|MMP1_HUMAN TIRGEVMFFKDRFYMRTNPFYPEVELNFISVFWPQLPNGLEAAYEFADRDEVRFFKGNKY 348
SP|P08254|MMP3_HUMAN TLRGEILIFKDRHFWRKSLRKLEPELHLISSFWPSLPSGVDAAYEVTSKDLVFIFKGNQF 360
                     *:***:::****.: *..    * **::** ***.**.*::****.:.:* * :****::

SP|P03956|MMP1_HUMAN WAVQGQNVLHGYPKDIYSSFGFPRTVKHIDAALSEENTGKTYFFVANKYWRYDEYKRSMD 408
SP|P08254|MMP3_HUMAN WAIRGNEVRAGYPRGIH-TLGFPPTVRKIDAAISDKEKNKTYFFVEDKYWRFDEKRNSME 419
                     **::*::*  ***:.*: ::*** **::****:*::::.****** :****:** :.**:

SP|P03956|MMP1_HUMAN PGYPKMIAHDFPGIGHKVDAVFMKDGFFYFFHGTRQYKFDPKTKRILTLQKANSWFNCRK 468
SP|P08254|MMP3_HUMAN PGFPKQIAEDFPGIDSKIDAVFEEFGFFYFFTGSSQLEFDPNAKKVTHTLKSNSWLNC-- 477
                     **:** **.*****. *:**** : ****** *: * :***::*::    *:***:**

SP|P03956|MMP1_HUMAN N 469
SP|P08254|MMP3_HUMAN -
```

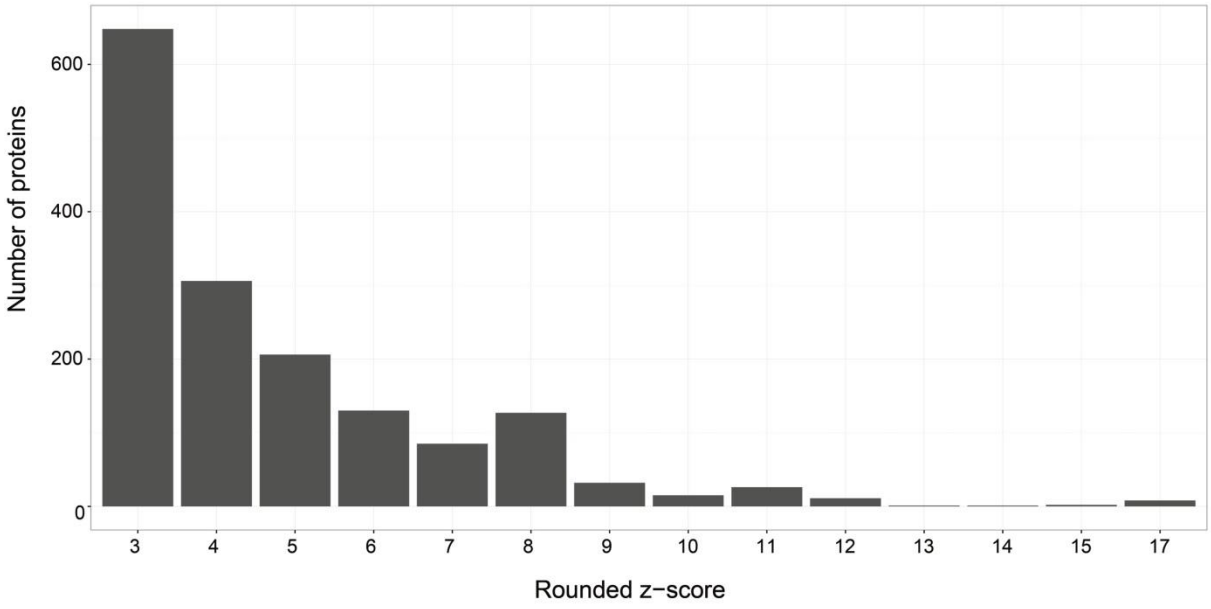
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 7G: Sequence homology search of MMP1-MBL2 using CLUSTALO alignment revealed a 10.33 % homology. CLUSTAL. In Yellow: Amino acidic sequence used as antigen in the generation of the antibody(PrEST).

```
CLUSTAL O(1.2.4) multiple sequence alignment

SP|P03956|MMP1_HUMAN MHSFPPLLLLLFWGVVSHSFPATLETQEQDVDLVQKYLEKYYNLKNDGRQVEKRRNSGPV 60
SP|P11226|MBL2_HUMAN MSLFPSLPLLL-LSMVAASYSETVTCEDAQKTCPA------------------------- 34
                     *  ** * ***  .:*: *:  *:  :: :

SP|P03956|MMP1_HUMAN VEKLKQMQEFFGLKVTGKPDAETLKVMKQPRCGVPDVAQFVLTEGNPRWEQTHLTYRIEN 120
SP|P11226|MBL2_HUMAN ----------------------------VIACSSPGINGFPGKDGRD------------- 53
                                                  *. *.:  *  .:*.

SP|P03956|MMP1_HUMAN YTPDLPRADVDHAIEKAFQLWSNVTPLTFTKVSEGQADIMISFVRGDHRDNSPFDGPGGN 180
SP|P11226|MBL2_HUMAN ----------------------------GTKGEKGEPG----------QGLRGLQGPPGK 75
                                                 ** .:*: .          :.   ::** *:

SP|P03956|MMP1_HUMAN LAHAFQPGP---------------GIGGDAHFDEDERWTNNFREYNLHRVAAHELGHSLG 225
SP|P11226|MBL2_HUMAN LGPPGNPGPSGSPGPKGQKGDPGKSPDGDSSLAASERKALQTEMARIKKWLTFSLGKQVG 135
                     *.   :***               . .**: :  .** : : .  .:::  :..**:.:*

SP|P03956|MMP1_HUMAN LSHSTDIGALMYPSYTFSGDVQLAQDDIDGIQAIYGRSQNPVQPIGPQTPKACDSKLTFD 285
SP|P11226|MBL2_HUMAN NKFFLTNGEIMT------------FEKVKALCVKF---QASV-----ATPRNAAEN---G 172
                      ..    * :*              :.:..: . :   *  *      **: . .:   .

SP|P03956|MMP1_HUMAN AITTIRGEVMFFKDRFYMRTNPFYPEVELNFISVFWPQLPNGLEAAYEFADRDEVRFFKG 345
SP|P11226|MBL2_HUMAN AIQNLIKEEAFLGITDEKTEGQFVDLTGNRLTYTNWNEGEP----NNAGSDEDCVLLL-K 227
                     ** .:  *  *:        . *   .  .:  . * :           :*.* * ::

SP|P03956|MMP1_HUMAN NKYWAVQGQNVLHGYPKDIYSSFGFPRTVKHIDAALSEENTGKTYFFVANKYWRYDEYKR 405
SP|P11226|MBL2_HUMAN NGQWNDVPCSTS------HLAVCEFPI-------------------------------- 248
                     *  *     ..         :   **

SP|P03956|MMP1_HUMAN SMDPGYPKMIAHDFPGIGHKVDAVFMKDGFFYFFHGTRQYKFDPKTKRILTLQKANSWFN 465
SP|P11226|MBL2_HUMAN ------------------------------------------------------------


SP|P03956|MMP1_HUMAN CRKN 469
SP|P11226|MBL2_HUMAN ----
```

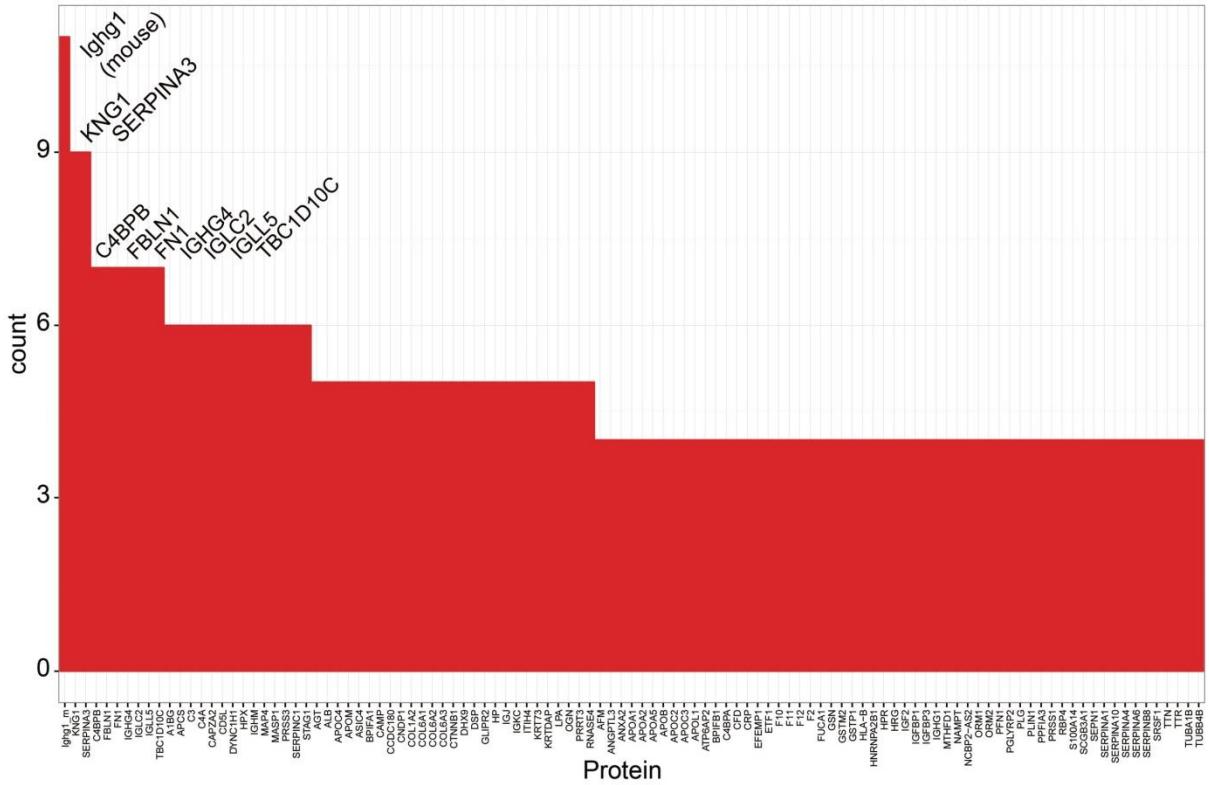
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 8. Bar plot representation of the number of proteins to which a z-score Ó3 was assigned. Values of z-scores were rounded.



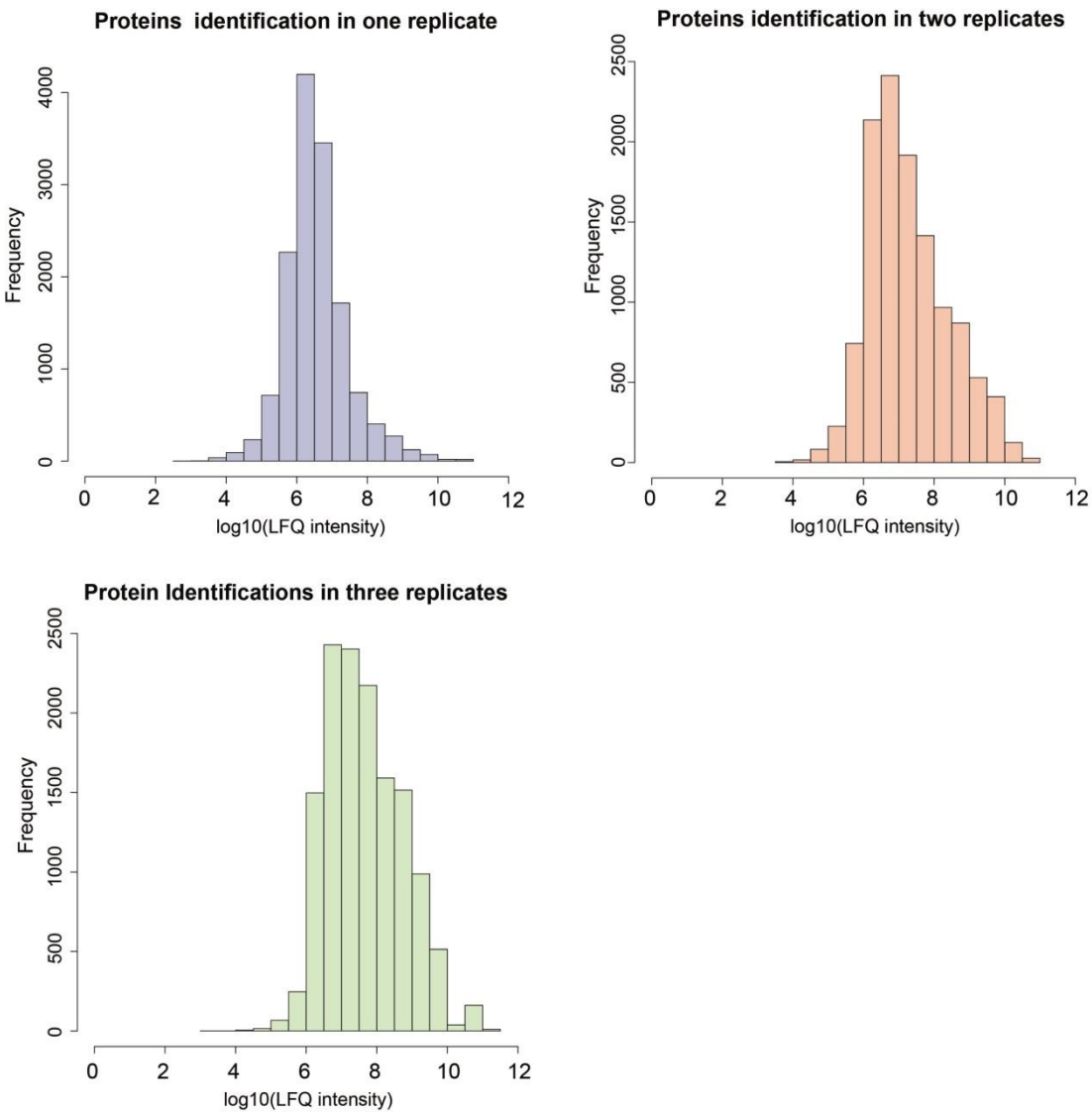
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 9: Bar plot representation of proteins identified as co-targets for more than 3 antibodies.

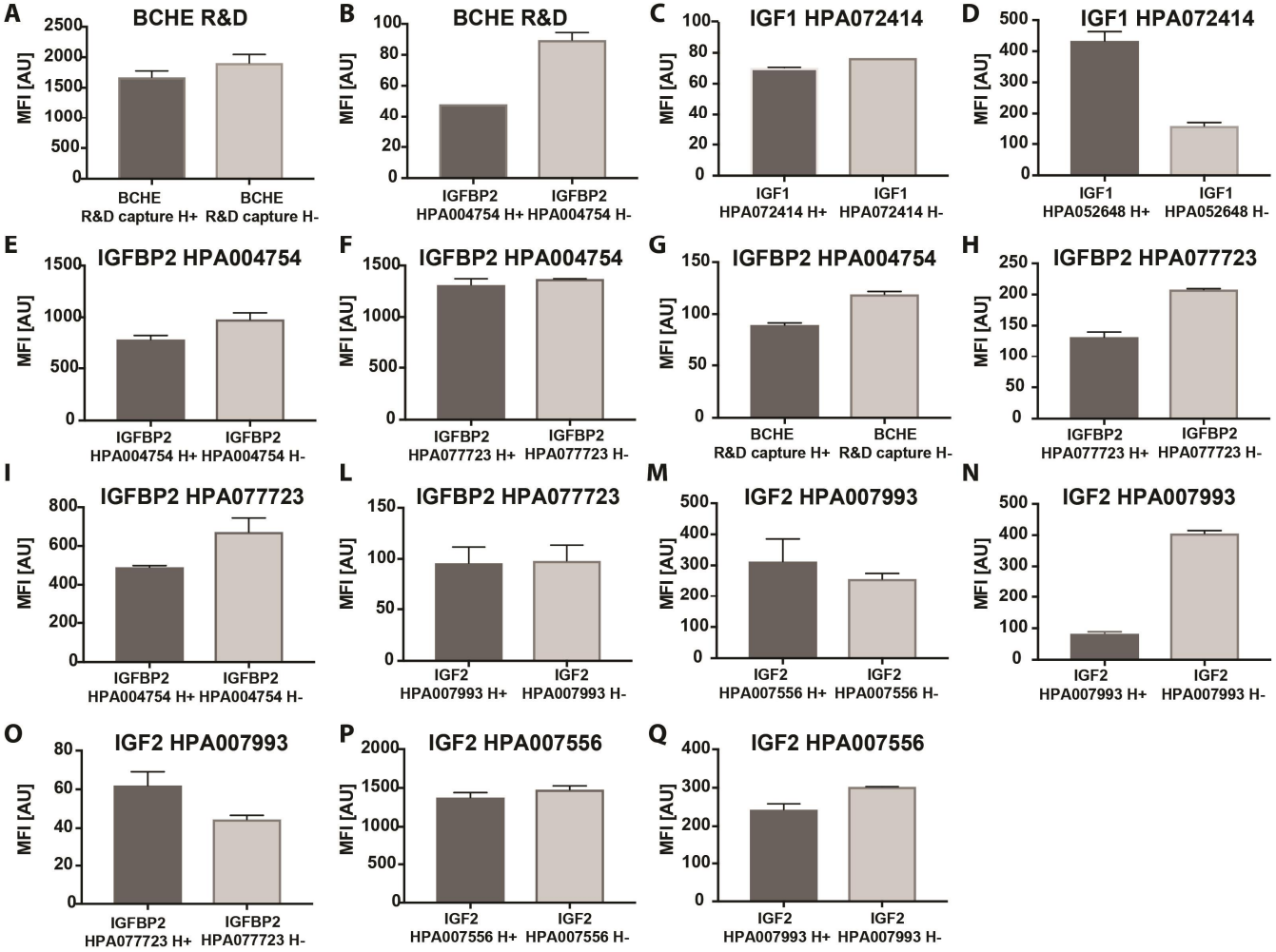


bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 10. Defining a LFQ cut-off base on the Distribution of LFQ intensities associated to proteins identified (A) in a single; (B) duplicate and (C) in triplicate experiments.



Supplementary Figure 11 Combinations of antibodies were tested in plasma. The MFI values of the corresponding pair of antibodies indicated for the EC50 plasma concentration. Headers of the plots indicate the detection antibody and on the x axis is the capture antibody. The data calculated is based on the mean values (±SD) of duplicate dilution curves. Dark grey (H+, heated plasma), Light grey (H-, untreated plasma).



bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Table 1. Comparison of validation rate by pIPs and WB for 104 antibodies assessed previously in plasma by WB. An antibody was considered passed by WB if scored 1,2 or 3, according to the Human Protein Atlas project.

	Plasma (MS)/Extracellular	**Cellular**	**Sum**
Method	N (%)	N (%)	N (%)
Pass by pIPs + WB*	11 (19)	2 (4)	13 (13)
Pass only by WB	8 (13)	13 (29)	21(20)
Pass only by pIPs	20(34)	12 (29)	32 (31)
WB (Total)	19 (32)	15 (33	
pIPs (Total)	32 (54)	15 (33)	
Failed by pIPs + WB*	20 (34)	18 (40)	38 (36)
Total*	59 (100)	45 (100)	104 (100)

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

SUPPLEMENTARY EXCEL TABLE

Description of content of each Sheet.

- ñSelected targetsò: Description of proteins for which antibodies were selected. (See Supplementary References 23-108)
- ñProtein Annotation (GO)ò: Protein annotation using GO terms.
- ñAntibodies experim. annotationsò: Antibodies used in the study, annotations about experimental conditions and data of validation by pIP and WB
- ñFrequencies of identificationò: List of proteins identified in pIP experiments and their frequencies of identification in heat treated or not heat treated plasma.
- ñBatches Kruskal Wallis Testò: p-values for differences in intensity levels detected the between batches
- ñGo enrichment analysisò: Differential GO analysis using TOPP cluster between frequent (> 20%) and less frequently occuring proteins (< 20%).
- ñz-score >2.5ò: List proteins for which a z-score > 2.5 was calculated containing targets, off-targets and potential interactors.
- ñList of Peptidesò: Peptides identified in pIP assays for the expected target proteins and overlapping with the antigen used to generate the antibodies.
- ñExperimental Batchesò: description of major differences in the experimental conditions for the 6 experimental batches analyzed.
- ñReagents Catalogs_Lot_numbersò: List of catalog and lot numbers for affinity reagents.
- ñAntibodies against same proteinò: proteins enriched by antibodies raised against the same target protein, using same or different protein fragments.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

MATERIAL AND METHODS

2.1 Sample collection.

Human EDTA plasma from a pool of individuals (50% females) was obtained from Seralab (Sera Laboratories International Ltd). Aliquots of plasma (0.5 mL) were stored in cryogenic vials at -80ºC and thawed at 4ºC before use.

2.2 Target selection

Information about target proteins their functions and involvement in diseases were collected through literature searches, Gene Ontology (GO, http://www.geneontology.org/), Human Protein Atlas (HPA, http://www.proteinatlas.org/), and Early Detection Research Network (EDRN, https://edrn.nci.nih.gov/). Plasma protein abundances were obtained from the 2017 built of the Plasma PeptideAtlas ([1], www.peptideatlas.org). Proteins were classified using the following terms: (i) Cellular or Extracellular, when the proteins appeared in one or more of the terms extracellular region (GO:0005576); extracellular space (GO:0005615); extracellular exosome (GO:0070062); proteinaceous extracellular matrix (GO:0005578); (see columns "GO CC Complete" and "Summary of GO CC"in the Excel sheet "Protein Annotation GO"). The list of antibodies and information was reported in Supplementary Excel Tables. The analysis included 157 antibodies. There were 15 monoclonal antibodies (10 from R&D Systems; 1 from HyTest Ltd.; 3 from Atlas Antibodies and 1 from SigmaAldrich) and144 polyclonal antibodies from the Human Protein Atlas [2]. In addition, normal rabbit IgG (Bethyl Laboratories), mouse IgG and rat IgG (both Santa Cruz Biotechnology) were included as controls. Catalog and lot numbers are listed in Supplementary Excel Tables. For antibodies obtained by the Human Protein Atlas, antibody IDs and lot numbers are the same.

2.3 Antibody coupling to magnetic beads.

Covalent coupling of antibody to magnetic beads (MagPlex, Luminex Corp.) was performed as previously described [3]. Briefly beads were activated using sulfo-NHS (Sulfo-N-hydroxysulfosuccinimide, Thermo) and ethyl-carbodiimide (EDC, both Thermo). Then 1,6 µg antibodies diluted in MES buffer per 500 000 beads were incubated 2 h at room temperature, beads were washed and stored in blocking buffer at 4 ºC.

2.4 Immunocapture-mass spectrometry

Aliquots of EDTA plasma (Seralab) were diluted 1:10 in assay buffer containing 0.5% w/v PVA (Sigma-Aldrich), 0.8% w/v PVP (Sigma-Aldrich), 0.1% w/v casein (Sigma-Aldrich) and 10 % of rabbit IgG (Bethyl Laboratories, Inc.). Samples undergoing heat treatment were incubated for 30 min at 56 ºC in water bath, before being combined with beads and incubated overnight on a rotation shaker at 23°C. For the final assessment of 153 antibodies 100 µl of crude plasma and 1,6 µg of antibody couple to beads were applied in each incubation. On the next day, and using a magnetic bead handler (KingFisher™ Flex Magnetic Particle Processors, Thermo Scientific), beads were separated from the

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

sample, washed with 0.03% Chaps in PBS and re-suspended in digestion buffer containing 50 mM ammonium bicarbonate (Sigma-Aldrich) and 0.25% sodium deoxycholate (Sigma-Aldrich). Proteins were reduced with 1 mM DTT (Sigma-Aldrich) at 56 ºC for 30 min, and alkylated by iodoacetamide 4 mM (Sigma-Aldrich), at RT in the dark for 30 min. Alkylation was quenched adding 1 mM DTT. Proteins were digested using a mixture of Trypsin and LysC at 1:25 trypsin-to-protein ratio (Promega, USA) overnight at 37 ºC. Enzyme inactivation and sodium deoxycholate precipitation was obtained adding 0.005% TFA. Peptides in the supernatant were then separated from beads, dried and re-suspended in solvent A containing 3% acetonitrile (ACN) and 0.1% formic acid (FA).

LC-MS/MS.

MS analysis was performed using a Q-Exactive HF (Thermo) operated in a data dependent mode, equipped with an Ultimate 3000 RSLC nanosystem, Dionex). Samples were injected into a C18 guard desalting column (Acclaim pepmap 100, 75 µm × 2 cm, nanoViper, P/N 164535, Thermo) and then into a 50 cm x 75µm ID Easy spray analytical column packed with 2µm C18 (EASY-Spray C18 P/N ES803, Thermo) for RPLC. Elution was performed in a linear gradient of Buffer B (90% ACN, 5% DMSO, 0.1% FA) from 3 % to 43% in 50 min at 250 nL/min. The proportion of Buffer B was increased stepwise to 45% in 5 min, then to 99% in 2 min, and then held for 10 minutes. DMSO was added Buffer A for the chromatography (90% water, 5% ACN, 5% DMSO, 0.1% FA). Full MS scan (300-1600 m/z) proceeded at resolution of 60,000. Precursors were isolated with a width of 2 m/z and listed for exclusion for 60 s. The top five most abundant ions were selected for higher energy collision dissociation (HCD). Single and unassigned charge states were rejected from precursor selection. In MS/MS, a max ion injection time of 250 ms and AGC target of 1E5 were applied.

2.7 Data analysis

Shotgun MS data search was performed on MaxQuant (v1.5.3.30) [4] using the integrated algorithm MaxLFQ. Spectra were search against a human protein database from Uniprot (accessed on 03/17/2016, Canonical and Isoforms, 20,198 hits customized adding sequences of immunoglobulins chain C from rabbit, rat and mouse, LysC (PSEAE) and Trypsin (PIG). Settings allowed for two missing cleavages, methionine oxidation and N-term acetylation as variable modification and cysteine carbamidomethylation as fixed modification. Fast LFQ and match between runs were applied, three minimum number of neighbors, and six average number of neighbors. All the 414 raw data files included in the analysis of 153 antibodies plus controls were analyzed in a single session, LFQ intensity values obtained were used for the following analysis. We considered as contaminants: proteins belonging to the list of contaminants in MaxQuant not belonging to Homo sapiens (Human), Ig gamma chain C region from Oryctolagus cuniculus (Rabbit)) because known to be in the dilution buffer andImmunoglobulin variable chains belonging to Homo sapiens (Human), and excluded from the z-score calculations. We considered missing values as missing not at random (MNAR) [5], but missing because of concentrations below the limit of detection (LOD). We therefore used min = 0 as

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

minimum value detected of intensity (Single-value imputation approach). When calculating average and standard deviation for each protein identified over the population of all experiments, missing values of LFQ intensities were substituted to 1 to allow for log10-transformation before two ways hierarchical clustering and principal component analysis. For z-scores calculation, when duplicate and triplicate experiments were available, we considered only proteins identified in all replicates for further analyses, and calculated average of LFQ intensities Proteins were considered enriched when associated to a z-score ≥ 3. To visualize the enriched proteins for each antibody, z-scores and LFQ intensity values were use and proteins found above the set threshold were annotated accordingly.
Raw data produced to assess experimental conditions were analyzed using MaxQuant but excluding the function for LFQ..

Data analysis and representation was performed on the environment for statistical computing and graphics R [6]. Alignments between protein and prEST sequences was performed using the Clustal Omega program available at EMBL-EBI [7]. GO enrichment system was performed using the PANTHER Classification System (http://pantherdb.org/). Comparison of GO terms was conducted using ToppCluster ([8]; https://toppcluster.cchmc.org/), regarding Bonferroni corrected p-values < 0.01 as significant.

2.8 Sandwich Immunoassay

The capture antibodies towards IGFBP2, IGF1, IGF2, DERA, BCHE, and rabbit-Immunoglobulin G (rIgG) and mouse-IgG, as negative controls (**Supplementary Excel Table**, sheet: "Reagents Lot numbers"), were covalently coupled to color-coded magnetic beads, A Suspension Bead Array (SBA) was generated and analyzed in Luminex Platform as previously described [9], using an in house protocol for labeling detection antibodies with biotin [10]. The antibodies were coupled to beads and labeled in order to test each different combination of capture and detection antibody pairs listed in the Supplementary Excel Table.
Plasma (EDTA Seralab, LOT#BRH1147432) was thawed on ice and centrifuged for 1 min at 2000 rpm, and diluted from 1:20 following 4-fold dilutions in PVX casein (PVXC) buffer 10% rIgG. The dilution series consisted of 6 points in duplicate and were heated at 56°C for 30 min. Then, the plasma was incubated with the SBA overnight. The same procedure was carried out with non-heated plasma dilution series.

The recombinant proteins used were IGFBP2 and IGF1 were a kind gift from Hanna Tegel and Johan Rockberg (AlbaNova University Center, KTH), IGF-II (R&D systems, catalog # 292-G2-050, lot DS2416011) and BCHE (DuoSet kit R&D systems, Catalog # DY6137-05, lot # 1387842). The dilution series of the proteins in buffer (PVXC 10% rIgG) consisted of 7 points prepared in duplicate and heated at 56°C for 30 min before incubation with the SBA overnight. Non-heated protein dilution series were also tested. The standard curves comprised different concentration ranges depending on

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the protein of interest. IGFBP2 and IGF2 were diluted in buffer from 500 ng/mL following 3-fold dilutions, IGF1 from 12000 pg/mL following 3-fold dilutions and BCHE 10000 pg/mL following 2-fold dilutions.

The detection antibodies were applied at 1µg/mL (HPA antibodies) or 25 ng/mL (BCHE R&D systems) for 90 min, and streptavidin- R-phycoerythrin (R-PE) conjugate (Life Technologies; SA10044) was used for the fluorescence read out in FlexMap3D (Luminex Corp.).

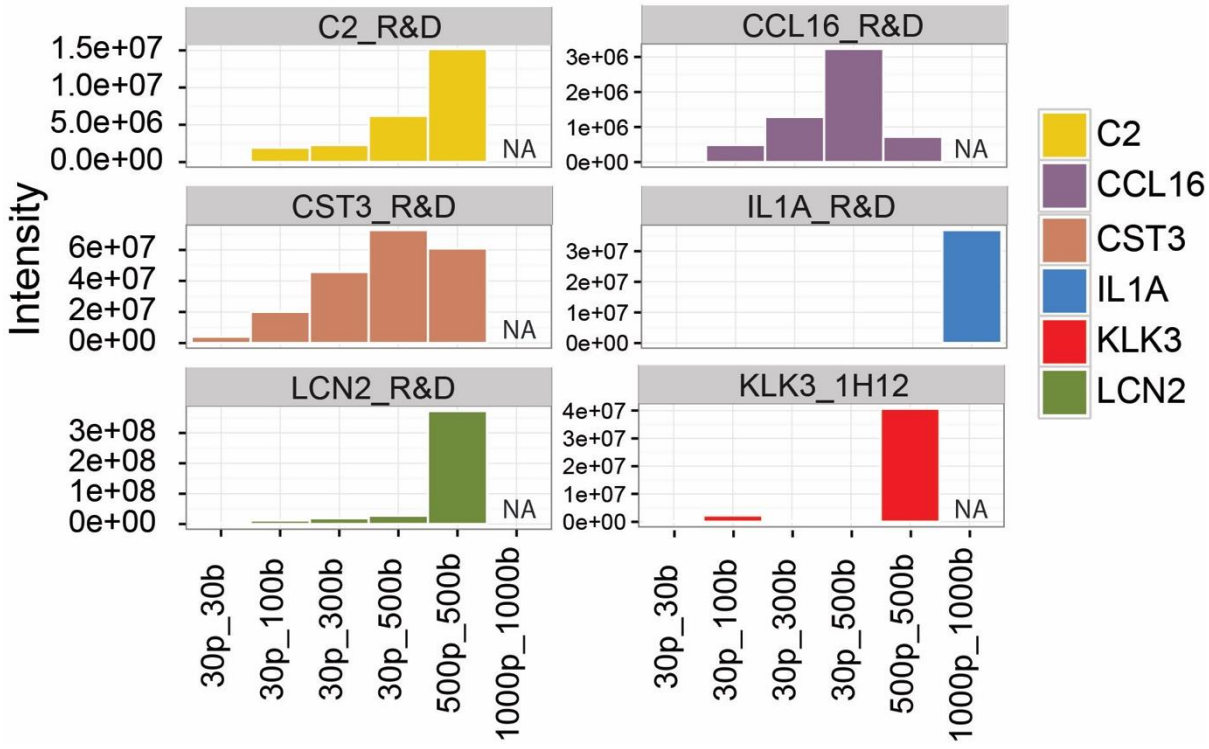
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

SUPPLEMENTARY NOTES

Supplementary note 1: Optimization of experimental conditions and quality control

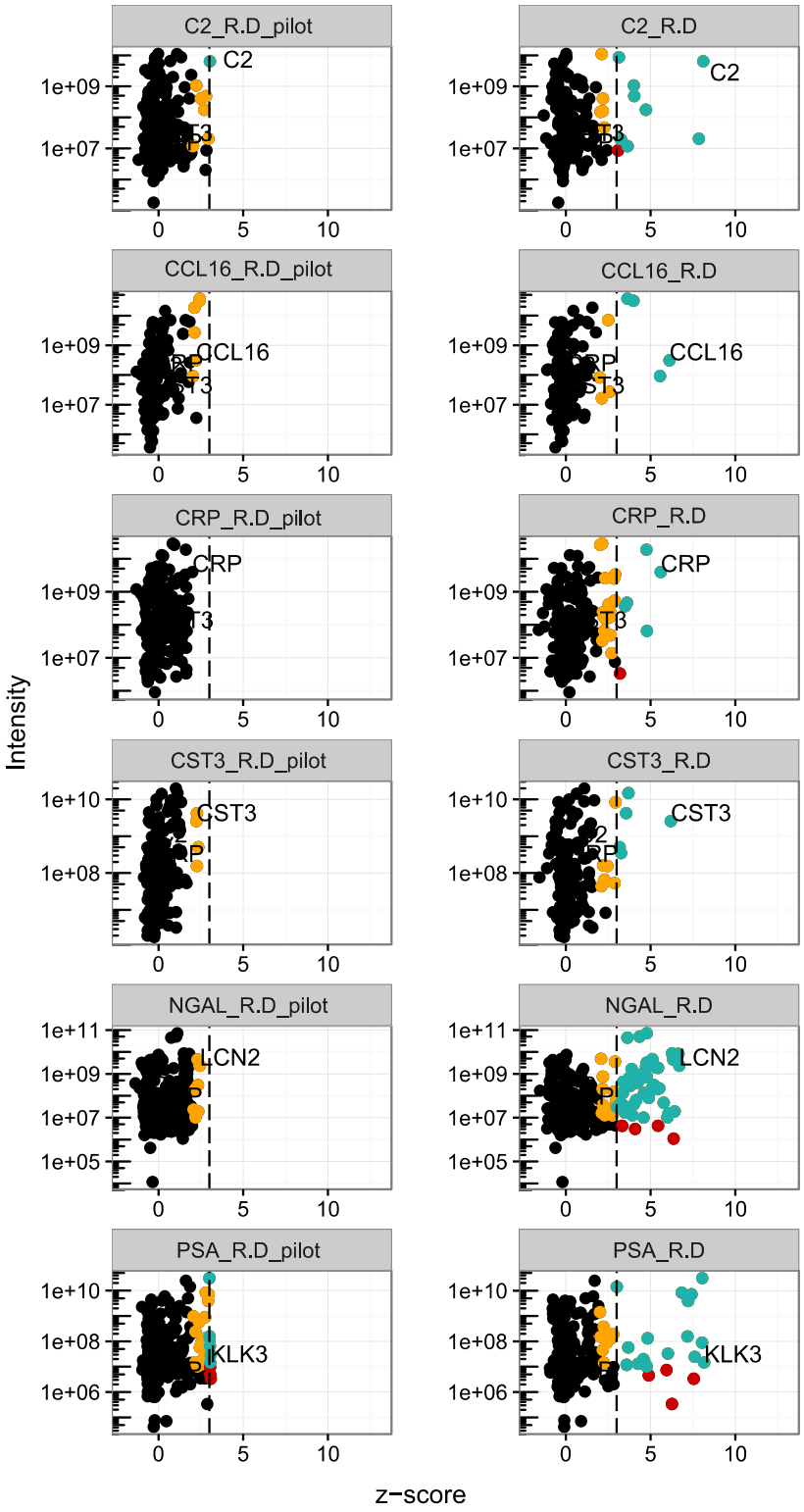
Factors such as target protein concentration and properties of ionization of the peptides of interest would ideally require optimized analysis for each antibody/target. Nevertheless, we began with developing a procedure applicable to a broad range of antibodies and target proteins. In order to set up optimal experimental conditions, we evaluated technical aspects such as volume of neat plasma andamount of antibody. As expected, increasing the volume of plasma, peptides belonging to less abundant proteins became detectable. For example interleukin 1 alpha (IL1A), with concentrations ~3 pg/mL in healthy human plasma [11], was detectable when 1 mL of plasma and 3.2 µg of antibody were applied (**Supplementary Figure 2A**). We established our protocol to enable the detection of proteins in a range of concentrations from µg/mL (C2) [12] to high pg/mL (KLK3) [13] [14-17] (**Supplementary Figure 2B**).

Supplementary Figure 2A: Optimization of experimental and technical conditions. Antibodies targeting known plasma proteins of different abundance (C2, CCL16, CST3, IL1A, KLK3, LCN2) were used to establish optimal experimental conditions for volume of plasma and number of beads coupled to antibody. In the axis legends, 'p' refers to the volume of plasma in µL; and 'b' to number of coupled beads (e.g. 500b = 500,000 beads = 1.6 µg of antibody).



bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 2B: Effect of library size. Data derived from IP assays with antibodies targeting the plasma proteins of C2, CCL16, CRP, CST3, NGAL and KLK3 (PSA) was analyzed using the pilot library limited to only these 21 IPs (left) as well as the library of more than 400 IP (right)**.** In this experiment, 100 µl of plasma and 500,000 beads were applied.



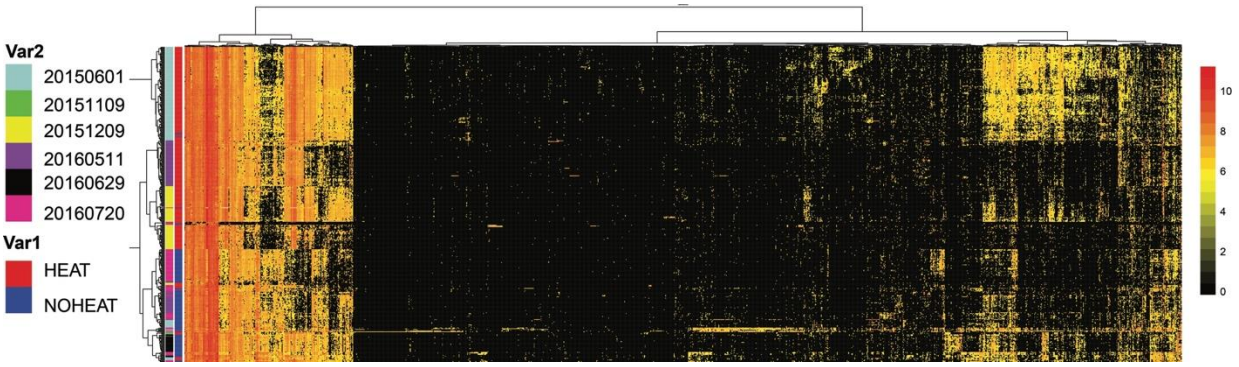
bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary note 2: Batch effects

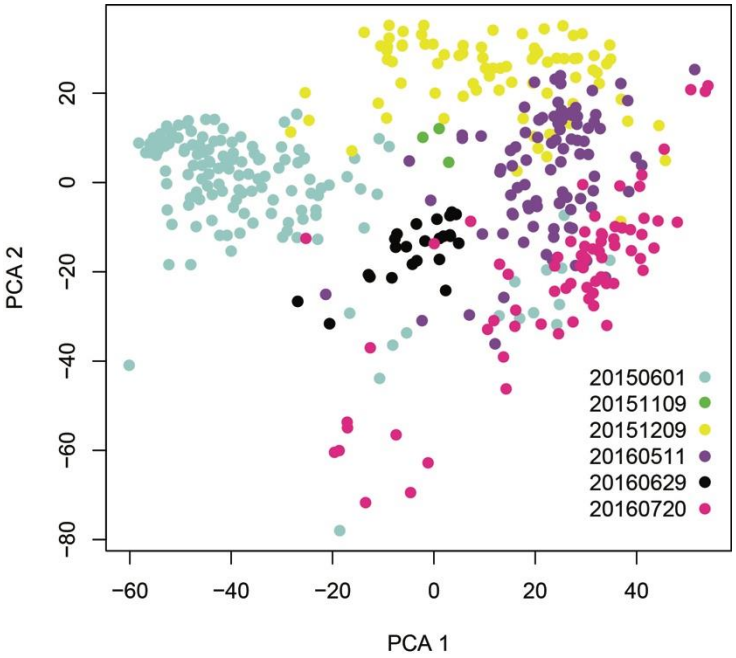
We observed that independent IPs performed in the same batch clustered together (**Supplementary Figure 3A**). Parameters varying between independent batches IPs include different lots of reference plasma, trypsin, and analytical columns. Long-term drift in instrumental response, sample handling and sample heat-treatment may also add additional variability. We found that the main difference between the different assays was due to heat treatment of the samples. Indeed, despite experimental batches, assays using either heat treated or untreated plasma clustered together (**Supplementary Figure 4A-B**). For this reason, we decide to analyze IPs with heat treated and not heat-treated plasma separately, in order to compare the enrichment profiles from the IPs with similar background.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 2A: Two ways hierarchical clustering analysis of the proteins identified in 414 immunoprecipitations. Clustering distance: "euclidean"; Clustering method: Ward. Cluster lanes left to the heat map represent (1) different experimental batches and (2) sample heat treatment. **(B)** Representation of principal component analysis, dots colors highlight separated experimental batches.



Supplementary Figure 2B: Representation of principal component analysis, dots colors highlight separated experimental batches.

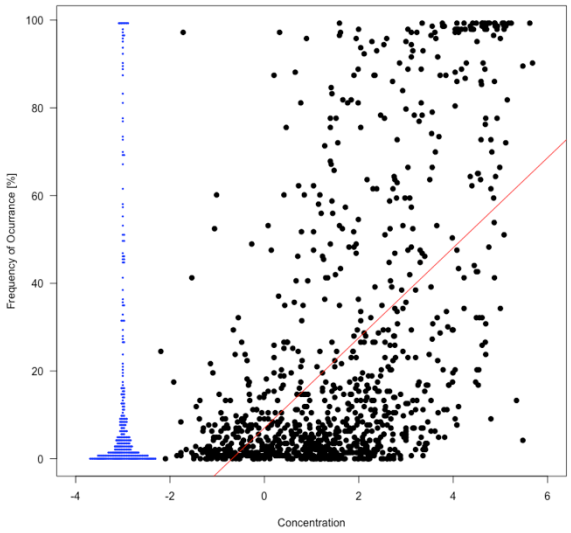


bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

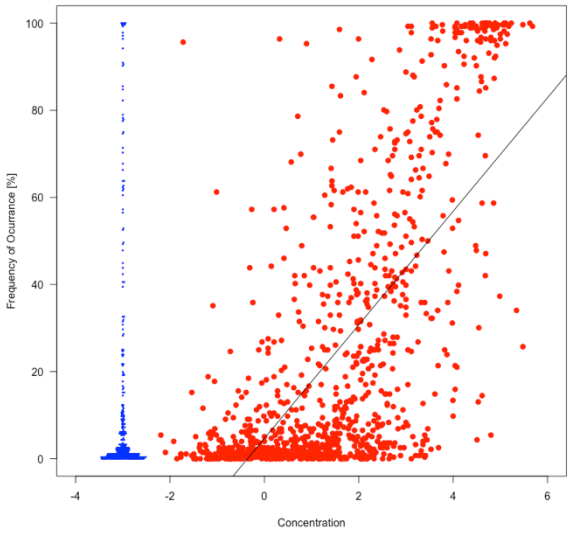
Supplementary note 3: Comparison of estimated plasma concentration and frequency

We investigated the relation between the frequency of occurrence of 1313 proteins in untreated plasma (**Supplementary Figure 3A**) and heat-treated plasma (**Supplementary Figure 3B**) with the estimated plasma concentration from PeptideAtlas. There was a significant association (p-value < 2.2e-16) for both sample types while 293/1313 = 22% of the proteins (blue dots) were not found with a concentration estimation in PeptideAtlas.

Supplementary Figure 3 A: Relation between estimated plasma concentration [ng/ml] and frequency [%] for untreated plasma



Supplementary Figure 3 B: Relation between estimated plasma concentration [ng/ml] and frequency [%] for heat-treated plasma.

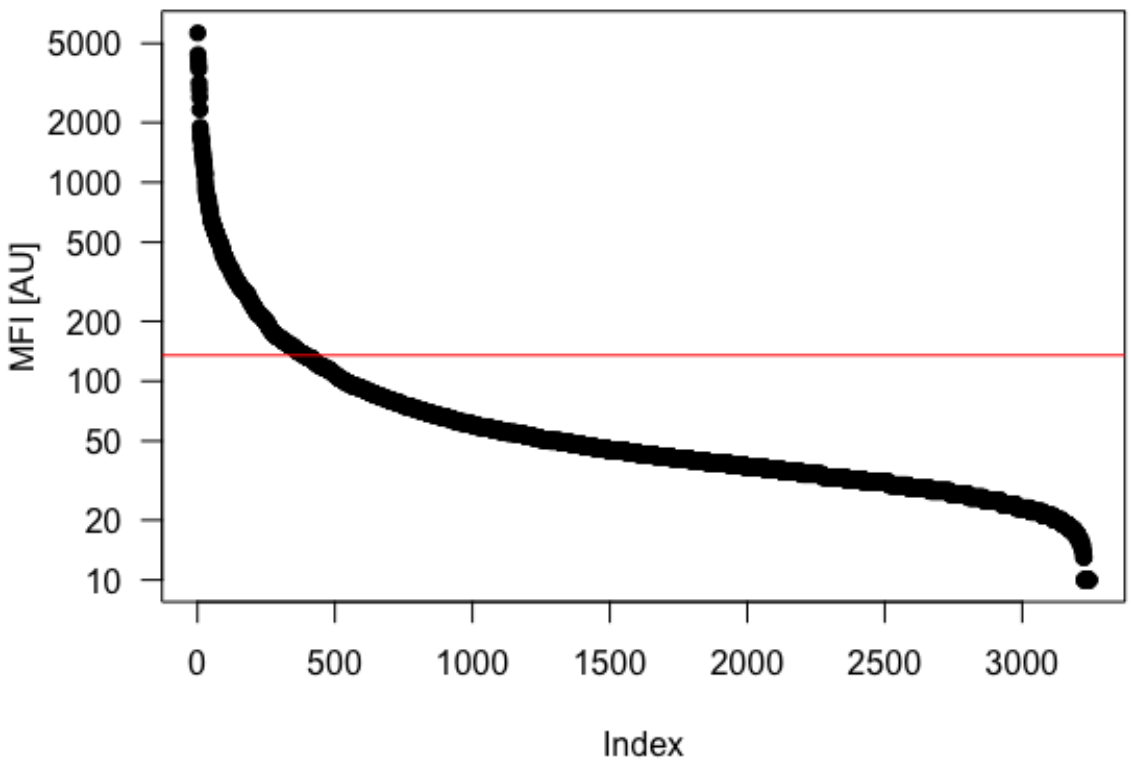


bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary note 4: Investigation of passenger proteins

We determined possible presence of antigens in the antibody solution. Antigens that were used during the affinity purification may have been co-eluted. Each antigen, denoted PrEST, consists of a tag (His_6 and albumin binding protein ABP) and the region selected target protein [18]. To detect the presence of the antigens, we chose an antibody specific to tag region (1:60,000; [19]) and incubated antibody-coupled beads accordingly. Detection using an anti-IgY antibody carrying R-PE (1 µg/ml) we determined the cut-off as 10x SD + mean MFI obtained from beads carrying normal rabbit IgG. Out of 4,180 antibodies tested (See Supplementary Figure 4), 41 had been investigated for IP-MS. Among these, 11 revealed MFI above cut-off (135 AU). These included on binder of the ON-target category (HPA042270), nine CO-target binders (HPA002869, HPA005905, HPA019493, HPA030603, HPA045140, HPA045822, HPA062231, HPA067590, HPA070841) and one OFF-target binder (HPA057179).

Supplementary Figure 6: Ranked distribution of the detection of passenger antigens on antibody coupled beads**.**



bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary note 5: Antibodies evaluation by protein arrays, WB and pIP.

In plasma proteomics, current options for assessing the selectivity of antibodies can be offered by paired antibodies, protein arrays [20] or Western blot (WB) [21]. Even if the three methods are each valuable in evaluating antibodies quality, validation of a previous discovery or qualification prior to a planned assay should be performed by a method resembling as much as possible the experimental conditions of the previous or intended application.

For both protein arrays and WB, the setup is that a surplus of antibodies is diluted in a solution and applied onto supports that present the antigens. Hence, there is generally minimal competition for binding sites between potential on- and off-targets as compared to antibodies being immobilized and applied to a complex solution. For an example given by anti-IL6R binder, we found application dependent recognition for five different antibodies. All five were classified as target-specific using protein arrays, however only three detected IL6R in plasma (**Supplementary Excel Table**). It is consequently the composition of plasma, where 90% of protein content is assigned to 20 proteins, that poses a challenge for WB in terms of analytical resolution. In samples in which the concentrations of certain proteins dominate, the efficiency of how proteins migrate through the gel can be affected. This can influence the distribution of proteins in terms of their molecular mass and abundance; hence limits the amount of plasma loaded on the gel and so detectability of less abundant targets.

As first, we compared the classifications obtained here with existing scores of validation assigned by plasma WB to the same antibody's Lot (HPA ID) within the Human Protein Atlas project [21]. Noteworthy, the comparison is based on different samples and the samples chosen for WB analysis were depleted of albumin and IgG prior use [22]. Nevertheless, we found that the assessment of 13 out of 104 antibodies (12%) provided supportive evidence by both methods. For antibodies raised against plasma proteins, the success rates for plasma IP (54%) was though higher than for plasma WB (32%) (**Supplementary Table 1**). When considering cellular proteins, the success rates were the same 33%. For WB however, uncertainty will remain unless other standards or comparative analyses are used to reference the detected bands. Until then, bands detected at the predicted molecular weight may still represent the recognition of an off-target molecule.

Consequently, pIP provides an unequivocal identification of the target and could elucidate ambiguous WB results (see C1orf64, CEP162, E2F7,and CCL16 **Supplementary Excel Table**, sheet: "Antibodies_experim_annotation"). While WB may provide a more accessible technology for some labs, the information added by identification in MS will be required for an in-depth analysis of antibody selectivity. IP-MS can here be one option to enhance our understanding of binding to proteins, in particular for plasma.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

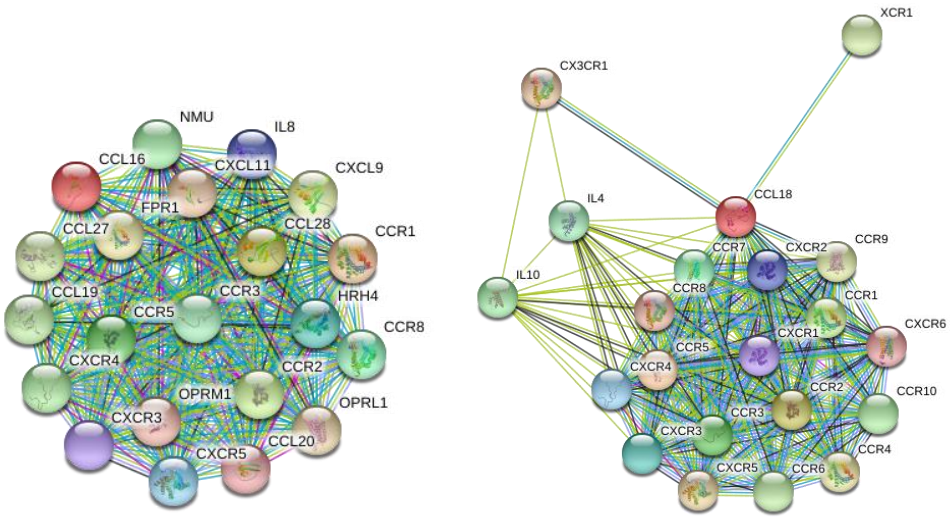
Supplementary Table 1. Comparison of validation rate by pIPs and WB for 104 antibodies assessed previously in plasma by WB. An antibody was considered passed by WB if scored 1,2 or 3, according to the Human Protein Atlas project.

	Plasma (MS)/Extracellular	**Cellular**	**Sum**
Method	N (%)	N (%)	N (%)
Pass by pIPs + WB*	11 (19)	2 (4)	13 (13)
Pass only by WB	8 (13)	13 (29)	20(20)
Pass only by pIPs	21(35)	13 (29)	34 (33)
WB (Total)	19 (32)	15 (33	
pIPs (Total)	32 (54)	15 (33)	
Failed by pIPs + WB*	19 (33)	17 (38)	36 (35)
Total*	59 (100)	45 (100)	104 (100)

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary note 7: Interaction and homology analysis

Supplementary Figure 7A: Interaction networks of CCL16 and CCL18 taken from String database V10.5 (20 interactors, 1st shell only).

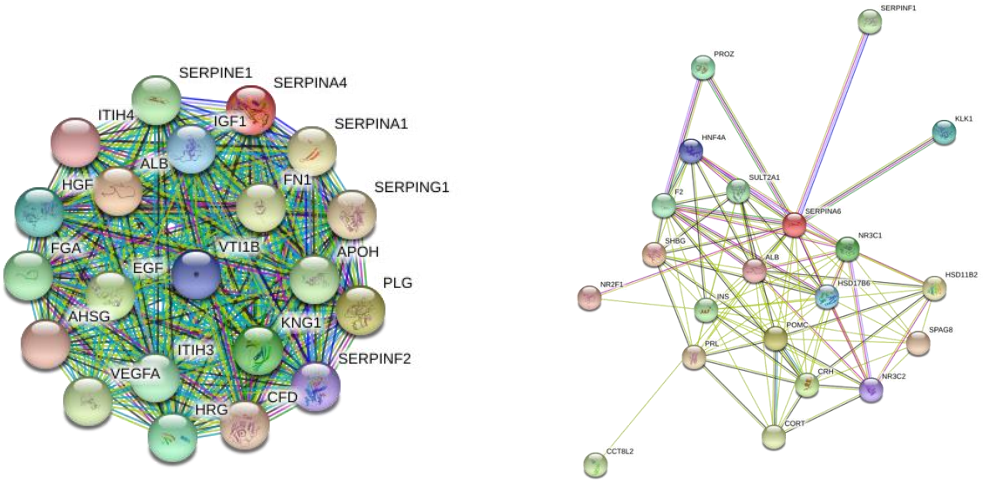


Supplementary Figure 7B: Sequence homology search of CCL16-CCL18using CLUSTALO alignment revealed a 26.6% homology. In Yellow: Amino acidic sequence used as antigen in the generation of the antibody.

```
CLUSTAL O(1.2.3) multiple sequence alignment

SP|O15467|CCL16_HUMAN MKVSEAALSLLVLILIITSASRSQPKVPEWVNTPSTCCLKYYEKVLPRRLVVGYRK-ALN 59
SP|P55774|CCL18_HUMAN MKGLAAA--LLVLVCTMALC-----SCAQVGTNKELCCLVYTSWQIPQKFIVDYSETSPQ 53
                      **   **  ****:  :: .     .  :  .. . *** * .  :*::::*.* : : :

SP|O15467|CCL16_HUMAN CHLPAIIFVTKRNREVCTNPNDDWVQEYIKDPNLPLLPTRNLSTVKIITAKNGQPQLLNS 119
SP|P55774|CCL18_HUMAN CPKPGVILLTKRGRQICADPNKKWVQKYISDLKLNA------------------------ 89
                      *  *.:*::***.*::*::**..***:**.* :*

SP|O15467|CCL16_HUMAN Q 120
SP|P55774|CCL18_HUMAN -
```

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 7C: Interaction networks of SERPINA4 and SERPINA6 taken from String database V10.5 (20 interactors, 1st shell only).

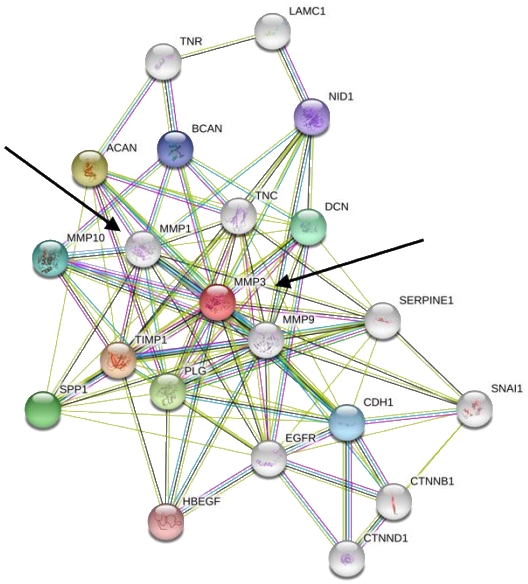


Supplementary Figure 7D: Sequence homology search of SERPINA4-SERPINA6 using CLUSTALO alignment revealed a 40.0% homology. In Yellow: Amino acidic sequence used as antigen in the generation of the antibody.

```
CLUSTAL O(1.2.3) multiple sequence alignment

SP|P29622|KAIN_HUMAN MHLIDYLLLLLVGLLALSHGQLHVEHDGESCSNSSHQQILETGEGSPSLKIAPANADFAF 60
SP|P08185|CBG_HUMAN  MPLLLYTCLLWLP----TSGLWTVQAMDP--------NAAYVNMSNHHRGLASANVDFAF 48
                     * *: *  ** :     : *   *:  .         :   .. ..    :* **.****

SP|P29622|KAIN_HUMAN RFYYLIASETPGKNIFFSPLSISAAYAMLSLGACSHSRSQILEGLGFNLTELSESDVHRG 120
SP|P08185|CBG_HUMAN  SLYKHLVALSPKKNIFISPVSISMALAMLSLGTCGHTRAQLLQGLGFNLTERSETEIHQG 108
                      :*  :.: :* ****:**:*** * ******:*.*:*:*:*:******** **:::*:*

SP|P29622|KAIN_HUMAN FQHLLHTLNLPGHGLETRVGSALFLSHNLKFLAKFLNDTMAVYEAKLFHTNFYDTVGTIQ 180
SP|P08185|CBG_HUMAN  FQHLHQLFAKSDTSLEMTMGNALFLDGSLELLESFSADIKHYYESEVLAMNFQDWATASR 168
                     **** : :   . .**  :*.****. .*::* .*  *    **::::  ** * . : :

SP|P29622|KAIN_HUMAN LINDHVKKETRGKIVDLVSELKKDVLMVLVNYIYFKALWEKPFISSRTTPKDFYVDENTT 240
SP|P08185|CBG_HUMAN  QINSYVKNKTQGKIVDLFSGLDSPAILVLVNYIFFKGTWTQPFDLASTREENFYVDETTV 228
                      **.:**::*:******.* *.. .::******:**. * :**  : *  ::*****.*.

SP|P29622|KAIN_HUMAN VRVPMMLQDQEHHWYLHDRYLPCSVLRMDYKGDATVFFILPNQGKMREIEEVLTPEMLMR 300
SP|P08185|CBG_HUMAN  VKVPMMLQSST-ISYLHDSELPCQLVQMNYVGNGTVFFILPDKGKMNTVIAALSRDTINR 287
                     *:******..    ****  ***.:::*:* *:.*******::***. :  .*: : : *

SP|P29622|KAIN_HUMAN WNNLLRKRNFYKKLELHLPKFSISGSYVLDQILPRLGFTDLFSKWADLSGITKQQKLEAS 360
SP|P08185|CBG_HUMAN  WSAGLTS----SQVDLYIPKVTISGVYDLGDVLEEMGIADLFTNQANFSRITQDAQLKSS 343
                     *.  * .    .:::*::**.:*** * *.::* .:*::***:: *::* **:: :*::*

SP|P29622|KAIN_HUMAN KSFHKATLDVDEAGTEAAAATSFAIKFFSAQTNRHILRFNRPFLVVIFSTSTQSVLFLGK 420
SP|P08185|CBG_HUMAN  KVVHKAVLQLNEEGVDTAGSTGVTLNL---TSKPIILRFNQPFIIMIFDHFTWSSLFLAR 400
                     * .***.*:::* *.::*.:*..::::    ::  *****:**:::**.  * * ***.:

SP|P29622|KAIN_HUMAN VVDPTKP 427
SP|P08185|CBG_HUMAN  VMNPV-- 405
                     *::*.
```

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Figure 7E: Interaction network of MMP1 and MMP3 taken from String database V10.5 (20 interactors, 1st shell only).



Supplementary Figure 7F: Sequence homology search of MMP1-MMP3 using CLUSTALO alignment revealed a 52.8 % homology. CLUSTAL. In Yellow: Amino acidic sequence used as antigen in the generation of the antibody.

```
CLUSTAL O(1.2.3) multiple sequence alignment

SP|P03956|MMP1_HUMAN MHSFPPLLLLLFWGVVSHSFPATLETQEQDVDLVQKYLEKYYNLKNDGRQVEKRRNSGPV 60
SP|P08254|MMP3_HUMAN MKSLPILLLLCVAVCSAYPLDGAARGEDTSMNLVQKYLENYYDLKKDVKQFVRRKDSGPV 60
                     *:*:* **** .    :: : .: . :: .::*******:**:**:* :*. :*::****

SP|P03956|MMP1_HUMAN VEKLKQMQEFFGLKVTGKPDAETLKVMKQPRCGVPDVAQFVLTEGNPRWEQTHLTYRIEN 120
SP|P08254|MMP3_HUMAN VKKIREMQKFLGLEVTGKLDSDTLEVMRKPRCGVPDVGHFRTFPGIPKWRKTHLTYRIVN 120
                     *:*:::**:*:**:**** *::**:**::********.:*    * *:*.:******* *

SP|P03956|MMP1_HUMAN YTPDLPRADVDHAIEKAFQLWSNVTPLTFTKVSEGQADIMISFVRGDHRDNSPFDGPGGN 180
SP|P08254|MMP3_HUMAN YTPDLPKDAVDSAVEKALKVWEEVTPLTFSRLYEGEADIMISFAVREHGDFYPFDGPGNV 180
                     ******:  ** *:***:::*.:******::: **:*******.  :* *  ******.

SP|P03956|MMP1_HUMAN LAHAFQPGPGIGGDAHFDEDERWTNNFREYNLHRVAAHELGHSLGLSHSTDIGALMYPSY 240
SP|P08254|MMP3_HUMAN LAHAYAPGPGINGDAHFDDDEQWTKDTTGTNLFLVAAHEIGHSLGLFHSANTEALMYPLY 240
                     ****: *****.******:**:**::    **. *****:****** **::  ***** *

SP|P03956|MMP1_HUMAN TFSGD---VQLAQDDIDGIQAIYGRSQNP---------VQPIGPQTPKACDSKLTFDAIT 288
SP|P08254|MMP3_HUMAN HSLTDLTRFRLSQDDINGIQSLYGPPPDSPETPLVPTEPVPPEPGTPANCDPALSFDAVS 300
                         *   .:*:****:***::**   :            *  * **  **  *:***::

SP|P03956|MMP1_HUMAN TIRGEVMFFKDRFYMRTNPFYPEVELNFISVFWPQLPNGLEAAYEFADRDEVRFFKGNKY 348
SP|P08254|MMP3_HUMAN TLRGEILIFKDRHFWRKSLRKLEPELHLISSFWPSLPSGVDAAYEVTSKDLVFIFKGNQF 360
                     *:***:::****.: *..    * **::** ***.**.*::****.:.:* * :****::

SP|P03956|MMP1_HUMAN WAVQGQNVLHGYPKDIYSSFGFPRTVKHIDAALSEENTGKTYFFVANKYWRYDEYKRSMD 408
SP|P08254|MMP3_HUMAN WAIRGNEVRAGYPRGIH-TLGFPPTVRKIDAAISDKEKNKTYFFVEDKYWRFDEKRNSME 419
                     **::*::*  ***:.*: ::*** **::****:*::::.****** :****:** :.**:

SP|P03956|MMP1_HUMAN PGYPKMIAHDFPGIGHKVDAVFMKDGFFYFFHGTRQYKFDPKTKRILTLQKANSWFNCRK 468
SP|P08254|MMP3_HUMAN PGFPKQIAEDFPGIDSKIDAVFEEFGFFYFFTGSSQLEFDPNAKKVTHTLKSNSWLNC-- 477
                     **:** **.*****. *:**** : ****** *: * :***::*::    *:***:**

SP|P03956|MMP1_HUMAN N 469
SP|P08254|MMP3_HUMAN -
```

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

SUPPLEMENTARY EXCEL TABLE

Description of content of each Sheet.

- "Selected targets": Description of proteins for which antibodies were selected. (See Supplementary References 23-108)
- "Protein Annotation (GO)": Protein annotation using GO terms.
- "Antibodies experim. annotations": Antibodies used in the study, annotations about experimental conditions and data of validation by pIP and WB
- "Frequencies of identification": List of proteins identified in pIP experiments and their frequencies of identification in heat treated or not heat treated plasma.
- "Batches Kruskal Wallis Test": p-values for differences in intensity levels detected the between batches
- "Go enrichment analysis": Differential GO analysis using TOPP cluster between frequent (> 20%) and less frequently occuring proteins (< 20%).
- "z-score >2.5": List proteins for which a z-score > 2.5 was calculated containing targets, off-targets and potential interactors.
- "List of Peptides": Peptides identified in pIP assays for the expected target proteins and overlapping with the antigen used to generate the antibodies.
- "Experimental Batches": description of major differences in the experimental conditions for the 6 experimental batches analyzed.
- "Reagents Catalogs_Lot_numbers": List of catalog and lot numbers for affinity reagents.
- "Antibodies against same protein": proteins enriched by antibodies raised against the same target protein, using same or different protein fragments.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

SUPPLEMENTARY REFERENCES

1. Schwenk JM, Omenn GS, Sun Z, Campbell DS, Baker MS, Overall CM, et al. The Human Plasma Proteome Draft of 2017: Building on the Human Plasma PeptideAtlas from Mass Spectrometry and Complementary Assays. J Proteome Res. 2017. doi: 10.1021/acs.jproteome.7b00467. PubMed PMID: 28938075.

2. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419. doi: 10.1126/science.1260419. PubMed PMID: 25613900.

3. Neiman M, Fredolini C, Johansson H, Lehtio J, Nygren PA, Uhlen M, et al. Selectivity analysis of single binder assays used in plasma protein profiling. Proteomics. 2013;13(23-24):3406-10. doi: 10.1002/pmic.201300030. PubMed PMID: 24151238; PubMed Central PMCID: PMCPMC4265267.

4. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008;26(12):1367-72. Epub 2008/11/26. doi: nbt.1511 [pii] 10.1038/nbt.1511. PubMed PMID: 19029910.

5. Lazar C, Gatto L, Ferro M, Bruley C, Burger T. Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies. J Proteome Res. 2016;15(4):1116-25. Epub 2016/03/01. doi: 10.1021/acs.jproteome.5b00981. PubMed PMID: 26906401.

6. Ihaka R, Gentleman R. R: a language for data analysis and graphics. Journal of Computional and Graphical Statistics. 1996;5:299-3214.

7. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539. Epub 2011/10/11. doi: 10.1038/msb.2011.75. PubMed PMID: 21988835; PubMed Central PMCID: PMCPMC3261699.

8. Kaimal V, Bardes EE, Tabar SC, Jegga AG, Aronow BJ. ToppCluster: a multiple gene list feature analyzer for comparative enrichment clustering and network-based dissection of biological systems. Nucleic Acids Res. 2010;38(Web Server issue):W96-102. doi: 10.1093/nar/gkq418. PubMed PMID: 20484371; PubMed Central PMCID: PMCPMC2896202.

9. Drobin K, Nilsson P, Schwenk JM. Highly multiplexed antibody suspension bead arrays for plasma protein profiling. Methods Mol Biol. 2013;1023:137-45. doi: 10.1007/978-1-4614-7209-4_8. PubMed PMID: 23765623.

10. Dezfouli M, Vickovic S, Iglesias MJ, Nilsson P, Schwenk JM, Ahmadian A. Magnetic bead assisted labeling of antibodies at nanogram scale. Proteomics. 2014;14(1):14-8. Epub 2013/12/04. doi: 10.1002/pmic.201300283. PubMed PMID: 24307663.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

11. Tamilselvi E, Haripriya D, Hemamalini M, Pushpa G, Swapna S. Association of disease severity with IL-1 levels in methotrexate-treated psoriasis patients. Scand J Immunol. 2013;78(6):545-53. doi: 10.1111/sji.12117. PubMed PMID: 24283773.

12. Borte S, von Döbeln U, Fasth A, Wang N, Janzi M, Winiarski J, et al. Neonatal screening for severe primary immunodeficiency diseases using high-throughput triplex real-time PCR. Blood. 2012;119(11):2552-5. Epub 2011/11/30. doi: 10.1182/blood-2011-08-371021. PubMed PMID: 22130802.

13. Lilja H, Ulmert D, Vickers AJ. Prostate-specific antigen and prostate cancer: prediction, detection and monitoring. Nat Rev Cancer. 2008;8(4):268-78. doi: 10.1038/nrc2351. PubMed PMID: 18337732.

14. Stowe RP, Peek MK, Cutchin MP, Goodwin JS. Plasma cytokine levels in a population-based study: relation to age and ethnicity. J Gerontol A Biol Sci Med Sci. 2010;65(4):429-33. Epub 2009/12/16. doi: 10.1093/gerona/glp198. PubMed PMID: 20018825; PubMed Central PMCID: PMCPMC2844059.

15. Del Valle-Pinero AY, Martino AC, Taylor TJ, Majors BL, Patel NS, Heitkemper MM, et al. Pro-inflammatory chemokine C-C motif ligand 16 (CCL-16) dysregulation in irritable bowel syndrome (IBS): a pilot study. Neurogastroenterology and motility : the official journal of the European Gastrointestinal Motility Society. 2011;23(12):1092-7. doi: 10.1111/j.1365-2982.2011.01792.x. PubMed PMID: 21951809; PubMed Central PMCID: PMC3557463.

16. Luc G, Bard JM, Lesueur C, Arveiler D, Evans A, Amouyel P, et al. Plasma cystatin-C and development of coronary heart disease: The PRIME Study. Atherosclerosis. 2006;185(2):375-80. Epub 2005/07/25. doi: 10.1016/j.atherosclerosis.2005.06.017. PubMed PMID: 16046222.

17. Murase K, Mori K, Yoshimura C, Aihara K, Chihara Y, Azuma M, et al. Association between plasma neutrophil gelatinase associated lipocalin level and obstructive sleep apnea or nocturnal intermittent hypoxia. PLoS One. 2013;8(1):e54184. doi: 10.1371/journal.pone.0054184. PubMed PMID: 23342100; PubMed Central PMCID: PMC3544801.

18. Nilsson P, Paavilainen L, Larsson K, Odling J, Sundberg M, Andersson AC, et al. Towards a human proteome atlas: high-throughput generation of mono-specific antibodies for tissue profiling. Proteomics. 2005;5(17):4327-37. PubMed PMID: 16237735.

19. Schwenk JM, Lindberg J, Sundberg M, Uhlen M, Nilsson P. Determination of binding specificities in highly multiplexed bead-based assays for antibody proteomics. Mol Cell Proteomics. 2007;6(1):125-32. doi: 10.1074/mcp.T600035-MCP200. PubMed PMID: 17060675.

20. Sjöberg R, Mattsson C, Andersson E, Hellström C, Uhlen M, Schwenk JM, et al. Exploration of high-density protein microarrays for antibody validation and autoimmunity profiling. N Biotechnol. 2016;33(5 Pt A):582-92. Epub 2015/09/28. doi: 10.1016/j.nbt.2015.09.002. PubMed PMID: 26417875.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

21. Algenas C, Agaton C, Fagerberg L, Asplund A, Bjorling L, Bjorling E, et al. Antibody performance in western blot applications is context-dependent. Biotechnol J. 2014;9(3):435-45. doi: 10.1002/biot.201300341. PubMed PMID: 24403002.

22. Eriksson C, Schwenk JM, Sjoberg A, Hober S. Affibody molecule-mediated depletion of HSA and IgG using different buffer compositions: a 15 min protocol for parallel processing of 1-48 samples. Biotechnol Appl Biochem. 2010;56(2):49-57. doi: 10.1042/BA20100041. PubMed PMID: 20446920.

23. Sosonkina N, Nakashima M, Ohta T, Niikawa N, Starenki D. Down-regulation of ABCC11 protein (MRP8) in human breast cancer. Exp Oncol. 2011;33(1):42-6. PubMed PMID: 21423094.

24. Khwairakpam AD, Shyamananda MS, Sailo BL, Rathnakaram SR, Padmavathi G, Kotoky J, et al. ATP citrate lyase (ACLY): a promising target for cancer prevention and treatment. Curr Drug Targets. 2015;16(2):156-63. PubMed PMID: 25537655.

25. Bjørklund SS, Kristensen VN, Seiler M, Kumar S, Alnæs GI, Ming Y, et al. Expression of an estrogen-regulated variant transcript of the peroxisomal branched chain fatty acid oxidase ACOX2 in breast carcinomas. BMC Cancer. 2015;15:524. Epub 2015/07/17. doi: 10.1186/s12885-015-1510-8. PubMed PMID: 26183823; PubMed Central PMCID: PMCPMC4504068.

26. Jacques J, Hotton D, Asselin A, De la Dure-Molla M, Coudert AE, Isaac J, et al. Ameloblastin as a putative marker of specific bone compartments. Connect Tissue Res. 2014;55 Suppl 1:117-20. doi: 10.3109/03008207.2014.923849. PubMed PMID: 25158194.

27. Liu Z, Lu H, Jiang Z, Pastuszyn A, Hu CA. Apolipoprotein l6, a novel proapoptotic Bcl-2 homology 3-only protein, induces mitochondria-mediated apoptosis in cancer cells. Mol Cancer Res. 2005;3(1):21-31. PubMed PMID: 15671246.

28. You J, Yang H, Lai Y, Simon L, Au J, Burkart AL. ARID2, p110α, p53, and β-catenin protein expression in hepatocellular carcinoma and clinicopathologic implications. Hum Pathol. 2015;46(7):1068-77. PubMed PMID: 26284269.

29. Huang RY, Su SG, Wu DC, Fu J, Zeng X. BLZF1 expression is of prognostic significance in hepatocellular carcinoma. Biochem Biophys Res Commun. 2015;467(3):602-9. Epub 2015/09/02. doi: 10.1016/j.bbrc.2015.08.119. PubMed PMID: 26342799.

30. Su D, Fu X, Fan S, Wu X, Wang XX, Fu L, et al. Role of ERRF, a novel ER-related nuclear factor, in the growth control of ER-positive human breast cancer cells. Am J Pathol. 2012;180(3):1189-201. doi: 10.1016/j.ajpath.2011.11.025. PubMed PMID: 22341523; PubMed Central PMCID: PMCPMC3349889.

31. Figueroa JE, Densen P. Infectious diseases associated with complement deficiencies. Clin Microbiol Rev. 1991;4(3):359-95. PubMed PMID: 1889047; PubMed Central PMCID: PMCPMC358203.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

32. Su DM, Zhang Q, Wang X, He P, Zhu YJ, Zhao J, et al. Two types of human malignant melanoma cell lines revealed by expression patterns of mitochondrial and survival-apoptosis genes: implications for malignant melanoma therapy. Mol Cancer Ther. 2009;8(5):1292-304. Epub 2009/04/21. doi: 10.1158/1535-7163.MCT-08-1030. PubMed PMID: 19383853; PubMed Central PMCID: PMCPMC3128982.

33. Tian GA, Zhu CC, Zhang XX, Zhu L, Yang XM, Jiang SH, et al. CCBE1 promotes GIST development through enhancing angiogenesis and mediating resistance to imatinib. Sci Rep. 2016;6:31071. Epub 2016/08/10. doi: 10.1038/srep31071. PubMed PMID: 27506146; PubMed Central PMCID: PMCPMC4978997.

34. Arakelyan A, Kriegova E, Kubistova Z, Mrazek F, Kverka M, du Bois RM, et al. Protein levels of CC chemokine ligand (CCL)15, CCL16 and macrophage stimulating protein in patients with sarcoidosis. Clin Exp Immunol. 2009;155(3):457-65. doi: 10.1111/j.1365-2249.2008.03832.x. PubMed PMID: 19220835; PubMed Central PMCID: PMCPMC2669522.

35. Shi L, Zhang B, Sun X, Zhang X, Lv S, Li H, et al. CC chemokine ligand 18(CCL18) promotes migration and invasion of lung cancer cells by binding to Nir1 through Nir1-ELMO1/DOC180 signaling pathway. Mol Carcinog. 2016;55(12):2051-62. Epub 2016/01/12. doi: 10.1002/mc.22450. PubMed PMID: 26756176.

36. Radulovich N, Pham NA, Strumpf D, Leung L, Xie W, Jurisica I, et al. Differential roles of cyclin D1 and D3 in pancreatic ductal adenocarcinoma. Mol Cancer. 2010;9:24. Epub 2010/02/01. doi: 10.1186/1476-4598-9-24. PubMed PMID: 20113529; PubMed Central PMCID: PMCPMC2824633.

37. Li Y, Qu P, Wu L, Li B, Du H, Yan C. Api6/AIM/Spα/CD5L overexpression in alveolar type II epithelial cells induces spontaneous lung adenocarcinoma. Cancer Res. 2011;71(16):5488-99. Epub 2011/06/22. doi: 10.1158/0008-5472.CAN-10-4225. PubMed PMID: 21697282; PubMed Central PMCID: PMCPMC3156382.

38. Awata T, Yamashita H, Kurihara S, Morita-Ohkubo T, Miyashita Y, Katayama S, et al. A genome-wide association study for diabetic retinopathy in a Japanese population: potential association with a long intergenic non-coding RNA. PLoS One. 2014;9(11):e111715. Epub 2014/11/03. doi: 10.1371/journal.pone.0111715. PubMed PMID: 25364816; PubMed Central PMCID: PMCPMC4218806.

39. Kalathas D, Theocharis DA, Bounias D, Kyriakopoulou D, Papageorgakopoulou N, Stavropoulos MS, et al. Chondroitin synthases I, II, III and chondroitin sulfate glucuronyltransferase expression in colorectal cancer. Mol Med Rep. 2011;4(2):363-8. Epub 2011/01/25. doi: 10.3892/mmr.2011.431. PubMed PMID: 21468578.

40. Bernardo BC, Belluoccio D, Rowley L, Little CB, Hansen U, Bateman JF. Cartilage intermediate layer protein 2 (CILP-2) is expressed in articular and meniscal cartilage and down-

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

regulated in experimental osteoarthritis. J Biol Chem. 2011;286(43):37758-67. Epub 2011/08/31. doi: 10.1074/jbc.M111.248039. PubMed PMID: 21880736; PubMed Central PMCID: PMCPMC3199518.

41. Shah SP, Morin RD, Khattra J, Prentice L, Pugh T, Burleigh A, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature. 2009;461(7265):809-13. doi: 10.1038/nature08489. PubMed PMID: 19812674.

42. Urbschat A, Obermüller N, Haferkamp A. Biomarkers of kidney injury. Biomarkers. 2011;16 Suppl 1:S22-30. doi: 10.3109/1354750X.2011.587129. PubMed PMID: 21707441.

43. Maxwell PJ, Neisen J, Messenger J, Waugh DJ. Tumor-derived CXCL8 signaling augments stroma-derived CCL2-promoted proliferation and CXCL12-mediated invasion of PTEN-deficient prostate cancer cells. Oncotarget. 2014;5(13):4895-908. doi: 10.18632/oncotarget.2052. PubMed PMID: 24970800; PubMed Central PMCID: PMCPMC4148108.

44. Melén E, Himes BE, Brehm JM, Boutaoui N, Klanderman BJ, Sylvia JS, et al. Analyses of shared genetic factors between asthma and obesity in children. J Allergy Clin Immunol. 2010;126(3):631-7.e1-8. doi: 10.1016/j.jaci.2010.06.030. PubMed PMID: 20816195; PubMed Central PMCID: PMCPMC2941152.

45. Chen WC, Chen MF, Lin PY. Significance of DNMT3b in oral cancer. PLoS One. 2014;9(3):e89956. Epub 2014/03/13. doi: 10.1371/journal.pone.0089956. PubMed PMID: 24625449; PubMed Central PMCID: PMCPMC3953114.

46. Barber AG, Castillo-Martin M, Bonal DM, Rybicki BA, Christiano AM, Cordon-Cardo C. Characterization of desmoglein expression in the normal prostatic gland. Desmoglein 2 is an independent prognostic factor for aggressive prostate cancer. PLoS One. 2014;9(6):e98786. Epub 2014/06/04. doi: 10.1371/journal.pone.0098786. PubMed PMID: 24896103; PubMed Central PMCID: PMCPMC4045811.

47. Chen X, Zhang Y, Shi Y, Lian H, Tu H, Han S, et al. MiR-129 triggers autophagic flux by regulating a novel Notch-1/ E2F7/Beclin-1 axis to impair the viability of human malignant glioma cells. Oncotarget. 2016;7(8):9222-35. doi: 10.18632/oncotarget.7003. PubMed PMID: 26824182; PubMed Central PMCID: PMCPMC4891036.

48. Chen J, Wei D, Zhao Y, Liu X, Zhang J. Overexpression of EFEMP1 correlates with tumor progression and poor prognosis in human ovarian carcinoma. PLoS One. 2013;8(11):e78783. Epub 2013/11/13. doi: 10.1371/journal.pone.0078783. PubMed PMID: 24236050; PubMed Central PMCID: PMCPMC3827232.

49. de Nigris F, Mega T, Berger N, Barone MV, Santoro M, Viglietto G, et al. Induction of ETS-1 and ETS-2 transcription factors is required for thyroid cell transformation. Cancer Res. 2001;61(5):2267-75. PubMed PMID: 11280797.

50. Komatsu H, Masuda T, Iguchi T, Nambara S, Sato K, Hu Q, et al. Clinical Significance of FANCD2 Gene Expression and its Association with Tumor Progression in Hepatocellular

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Carcinoma. Anticancer Res. 2017;37(3):1083-90. doi: 10.21873/anticanres.11420. PubMed PMID: 28314268.

51. Cheng M, Watson PH, Paterson JA, Seidah N, Chretien M, Shiu RP. Pro-protein convertase gene expression in human breast cancer. Int J Cancer. 1997;71(6):966-71. PubMed PMID: 9185698.

52. Ertao Z, Jianhui C, Chuangqi C, Changjiang Q, Sile C, Yulong H, et al. Low level of FOXL1 indicates a worse prognosis for gastric cancer patients. Tumour Biol. 2016;37(8):11331-7. Epub 2016/03/09. doi: 10.1007/s13277-016-4890-8. PubMed PMID: 26960689.

53. Ranta N, Turpeinen H, Oksanen A, Hämäläinen S, Huttunen R, Uusitalo-Seppälä R, et al. The Plasma Level of Proprotein Convertase FURIN in Patients with Suspected Infection in the Emergency Room: A Prospective Cohort Study. Scand J Immunol. 2015;82(6):539-46. doi: 10.1111/sji.12386. PubMed PMID: 26346780.

54. Ko SY, Naora H. HOXA9 promotes homotypic and heterotypic cell interactions that facilitate ovarian cancer dissemination via its induction of P-cadherin. Mol Cancer. 2014;13:170. Epub 2014/07/14. doi: 10.1186/1476-4598-13-170. PubMed PMID: 25023983; PubMed Central PMCID: PMCPMC4105245.

55. Denduluri SK, Idowu O, Wang Z, Liao Z, Yan Z, Mohammed MK, et al. Insulin-like growth factor (IGF) signaling in tumorigenesis and the development of cancer drug resistance. Genes Dis. 2015;2(1):13-25. doi: 10.1016/j.gendis.2014.10.004. PubMed PMID: 25984556; PubMed Central PMCID: PMCPMC4431759.

56. Cao Y, Nimptsch K, Shui IM, Platz EA, Wu K, Pollak MN, et al. Prediagnostic plasma IGFBP-1, IGF-1 and risk of prostate cancer. Int J Cancer. 2015;136(10):2418-26. Epub 2014/11/10. doi: 10.1002/ijc.29295. PubMed PMID: 25348852; PubMed Central PMCID: PMCPMC4360136.

57. Yau SW, Azar WJ, Sabin MA, Werther GA, Russo VC. IGFBP-2 - taking the lead in growth, metabolism and cancer. J Cell Commun Signal. 2015;9(2):125-42. Epub 2015/01/25. doi: 10.1007/s12079-015-0261-2. PubMed PMID: 25617050; PubMed Central PMCID: PMCPMC4458250.

58. Alizadeh AA, Bohen SP, Lossos C, Martinez-Climent JA, Ramos JC, Cubedo-Gil E, et al. Expression profiles of adult T-cell leukemia-lymphoma and associations with clinical responses to zidovudine and interferon alpha. Leuk Lymphoma. 2010;51(7):1200-16. doi: 10.3109/10428191003728628. PubMed PMID: 20370541; PubMed Central PMCID: PMCPMC4296320.

59. Bankaitis KV, Fingleton B. Targeting IL4/IL4R for the treatment of epithelial cancer metastasis. Clin Exp Metastasis. 2015;32(8):847-56. Epub 2015/09/18. doi: 10.1007/s10585-015-9747-9. PubMed PMID: 26385103; PubMed Central PMCID: PMCPMC4651701.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

60. Kumari N, Dwarakanath BS, Das A, Bhatt AN. Role of interleukin-6 in cancer progression and therapeutic resistance. Tumour Biol. 2016;37(9):11553-72. Epub 2016/06/03. doi: 10.1007/s13277-016-5098-7. PubMed PMID: 27260630.

61. Miller MA, Sullivan RJ, Lauffenburger DA. Molecular Pathways: Receptor Ectodomain Shedding in Treatment, Resistance, and Monitoring of Cancer. Clin Cancer Res. 2017;23(3):623-9. Epub 2016/11/28. doi: 10.1158/1078-0432.CCR-16-0869. PubMed PMID: 27895032; PubMed Central PMCID: PMCPMC5290119.

62. Garaud S, Willard-Gallo K. IRF5: a rheostat for tumor-infiltrating lymphocyte trafficking in breast cancer? Immunol Cell Biol. 2015;93(5):425-6. doi: 10.1038/icb.2015.39. PubMed PMID: 26010613.

63. Moore KM, Thomas GJ, Duffy SW, Warwick J, Gabe R, Chou P, et al. Therapeutic targeting of integrin αvβ6 in breast cancer. J Natl Cancer Inst. 2014;106(8). Epub 2014/06/28. doi: 10.1093/jnci/dju169. PubMed PMID: 24974129; PubMed Central PMCID: PMCPMC4151855.

64. Coustan-Smith E, Song G, Clark C, Key L, Liu P, Mehrpooya M, et al. New markers for minimal residual disease detection in acute lymphoblastic leukemia. Blood. 2011;117(23):6267-76. Epub 2011/04/12. doi: 10.1182/blood-2010-12-324004. PubMed PMID: 21487112; PubMed Central PMCID: PMCPMC3122946.

65. Ferreyra Solari NE, Belforte FS, Canedo L, Videla-Richardson GA, Espinosa JM, Rossi M, et al. The NSL Chromatin-Modifying Complex Subunit KANSL2 Regulates Cancer Stem-like Properties in Glioblastoma That Contribute to Tumorigenesis. Cancer Res. 2016;76(18):5383-94. Epub 2016/07/12. doi: 10.1158/0008-5472.CAN-15-3159. PubMed PMID: 27406830; PubMed Central PMCID: PMCPMC5026635.

66. Abella V, Scotece M, Conde J, Gómez R, Lois A, Pino J, et al. The potential of lipocalin-2/NGAL as biomarker for inflammatory and metabolic diseases. Biomarkers. 2015;20(8):565-71. Epub 2015/12/15. doi: 10.3109/1354750X.2015.1123354. PubMed PMID: 26671823; PubMed Central PMCID: PMCPMC4819811.

67. Koide N, Kasamatsu A, Endo-Sakamoto Y, Ishida S, Shimizu T, Kimura Y, et al. Evidence for Critical Role of Lymphocyte Cytosolic Protein 1 in Oral Cancer. Sci Rep. 2017;7:43379. Epub 2017/02/23. doi: 10.1038/srep43379. PubMed PMID: 28230172; PubMed Central PMCID: PMCPMC5322526.

68. Manier S, Powers JT, Sacco A, Glavey SV, Huynh D, Reagan MR, et al. The LIN28B/let-7 axis is a novel therapeutic pathway in multiple myeloma. Leukemia. 2017;31(4):853-60. Epub 2016/10/24. doi: 10.1038/leu.2016.296. PubMed PMID: 27773931; PubMed Central PMCID: PMCPMC5382134.

69. Wang Y, Shan Q, Hou G, Zhang J, Bai J, Lv X, et al. Discovery of potential colorectal cancer serum biomarkers through quantitative proteomics on the colonic tissue interstitial fluids from

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the AOM-DSS mouse model. J Proteomics. 2016;132:31-40. Epub 2015/11/12. doi: 10.1016/j.jprot.2015.11.013. PubMed PMID: 26581642.

70. Radon TP, Massat NJ, Jones R, Alrawashdeh W, Dumartin L, Ennis D, et al. Identification of a Three-Biomarker Panel in Urine for Early Detection of Pancreatic Adenocarcinoma. Clin Cancer Res. 2015;21(15):3512-21. doi: 10.1158/1078-0432.CCR-14-2467. PubMed PMID: 26240291; PubMed Central PMCID: PMCPMC4539580.

71. Rosas IO, Richards TJ, Konishi K, Zhang Y, Gibson K, Lokshin AE, et al. MMP1 and MMP7 as potential peripheral blood biomarkers in idiopathic pulmonary fibrosis. PLoS Med. 2008;5(4):e93. doi: 10.1371/journal.pmed.0050093. PubMed PMID: 18447576; PubMed Central PMCID: PMCPMC2346504.

72. Liguori L, Andolfo I, de Antonellis P, Aglio V, di Dato V, Marino N, et al. The metallophosphodiesterase Mpped2 impairs tumorigenesis in neuroblastoma. Cell Cycle. 2012;11(3):569-81. Epub 2012/02/01. doi: 10.4161/cc.11.3.19063. PubMed PMID: 22262177.

73. Sakata-Yanagimoto M, Chiba S. Notch2 and immune function. Curr Top Microbiol Immunol. 2012;360:151-61. doi: 10.1007/82_2012_235. PubMed PMID: 22695918.

74. Grandclement C, Borg C. Neuropilins: a new target for cancer therapy. Cancers (Basel). 2011;3(2):1899-928. Epub 2011/04/08. doi: 10.3390/cancers3021899. PubMed PMID: 24212788; PubMed Central PMCID: PMCPMC3757396.

75. Lee DI, Zhu G, Sasaki T, Cho GS, Hamdani N, Holewinski R, et al. Phosphodiesterase 9A controls nitric-oxide-independent cGMP and hypertrophic heart disease. Nature. 2015;519(7544):472-6. Epub 2015/03/18. doi: 10.1038/nature14332. PubMed PMID: 25799991; PubMed Central PMCID: PMCPMC4376609.

76. Ikenaga EH, Talib LL, Ferreira AS, Machado-Vieira R, Forlenza OV, Gattaz WF. Reduced activities of phospholipases A2 in platelets of drug-naïve bipolar disorder patients. Bipolar Disord. 2015;17(1):97-101. Epub 2014/07/08. doi: 10.1111/bdi.12229. PubMed PMID: 25041493.

77. Mealer RG, Murray AJ, Shahani N, Subramaniam S, Snyder SH. Rhes, a striatal-selective protein implicated in Huntington disease, binds beclin-1 and activates autophagy. J Biol Chem. 2014;289(6):3547-54. Epub 2013/12/09. doi: 10.1074/jbc.M113.536912. PubMed PMID: 24324270; PubMed Central PMCID: PMCPMC3916556.

78. Grawenda AM, O'Neill E. Clinical utility of RASSF1A methylation in human malignancies. Br J Cancer. 2015;113(3):372-81. Epub 2015/07/09. doi: 10.1038/bjc.2015.221. PubMed PMID: 26158424; PubMed Central PMCID: PMCPMC4522630.

79. Yu J, Liang Q, Wang J, Wang K, Gao J, Zhang J, et al. REC8 functions as a tumor suppressor and is epigenetically downregulated in gastric cancer, especially in EBV-positive subtype. Oncogene. 2017;36(2):182-93. Epub 2016/05/23. doi: 10.1038/onc.2016.187. PubMed PMID: 27212034; PubMed Central PMCID: PMCPMC5241426.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

80. Bystrom S, Fredolini C, Edqvist PH, Nyaiesh EN, Drobin K, Uhlen M, et al. Affinity Proteomics Exploration of Melanoma Identifies Proteins in Serum with Associations to T-Stage and Recurrence. Translational oncology. 2017;10(3):385-95. doi: 10.1016/j.tranon.2017.03.002. PubMed PMID: 28433799; PubMed Central PMCID: PMCPMC5403766.

81. Strand AD, Aragaki AK, Shaw D, Bird T, Holton J, Turner C, et al. Gene expression in Huntington's disease skeletal muscle: a potential biomarker. Hum Mol Genet. 2005;14(13):1863-76. Epub 2005/05/11. doi: 10.1093/hmg/ddi192. PubMed PMID: 15888475.

82. Swart JF, de Roock S, Prakken BJ. Understanding inflammation in juvenile idiopathic arthritis: How immune biomarkers guide clinical strategies in the systemic onset subtype. Eur J Immunol. 2016;46(9):2068-77. doi: 10.1002/eji.201546092. PubMed PMID: 27461267.

83. Donato R, Sorci G, Giambanco I. S100A6 protein: functional roles. Cell Mol Life Sci. 2017. Epub 2017/04/17. doi: 10.1007/s00018-017-2526-9. PubMed PMID: 28417162.

84. Lines KE, Chelala C, Dmitrovic B, Wijesuriya N, Kocher HM, Marshall JF, et al. S100P-binding protein, S100PBP, mediates adhesion through regulation of cathepsin Z in pancreatic cancer cells. Am J Pathol. 2012;180(4):1485-94. Epub 2012/02/11. doi: 10.1016/j.ajpath.2011.12.031. PubMed PMID: 22330678.

85. Spiegel S, Milstien S. The outs and the ins of sphingosine-1-phosphate in immunity. Nat Rev Immunol. 2011;11(6):403-15. Epub 2011/05/06. doi: 10.1038/nri2974. PubMed PMID: 21546914; PubMed Central PMCID: PMCPMC3368251.

86. Richard S, Lagerstedt L, Burkhard PR, Debouverie M, Turck N, Sanchez JC. E-selectin and vascular cell adhesion molecule-1 as biomarkers of 3-month outcome in cerebrovascular diseases. J Inflamm (Lond). 2015;12:61. Epub 2015/11/04. doi: 10.1186/s12950-015-0106-z. PubMed PMID: 26543408; PubMed Central PMCID: PMCPMC4634720.

87. Kumar P, Nandi S, Tan TZ, Ler SG, Chia KS, Lim WY, et al. Highly sensitive and specific novel biomarkers for the diagnosis of transitional bladder carcinoma. Oncotarget. 2015;6(15):13539-49. doi: 10.18632/oncotarget.3841. PubMed PMID: 25915536; PubMed Central PMCID: PMCPMC4537032.

88. Kim YI, Ahn JM, Sung HJ, Na SS, Hwang J, Kim Y, et al. Meta-markers for the differential diagnosis of lung cancer and lung disease. J Proteomics. 2016;148:36-43. Epub 2016/05/07. doi: 10.1016/j.jprot.2016.04.052. PubMed PMID: 27168012.

89. Newbold RF, Mokbel K. Evidence for a tumour suppressor function of SETD2 in human breast cancer: a new hypothesis. Anticancer Res. 2010;30(9):3309-11. PubMed PMID: 20944102.

90. Hsiao JH, Fu Y, Hill AF, Halliday GM, Kim WS. Elevation in sphingomyelin synthase activity is associated with increases in amyloid-beta peptide generation. PLoS One. 2013;8(8):e74016. Epub 2013/08/20. doi: 10.1371/journal.pone.0074016. PubMed PMID: 23977395; PubMed Central PMCID: PMCPMC3748018.

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

91. Ahn R, Sabourin V, Bolt AM, Hébert S, Totten S, De Jay N, et al. The Shc1 adaptor simultaneously balances Stat1 and Stat3 activity to promote breast cancer immune suppression. Nat Commun. 2017;8:14638. Epub 2017/03/09. doi: 10.1038/ncomms14638. PubMed PMID: 28276425; PubMed Central PMCID: PMCPMC5347092.

92. Urtasun R, Elizalde M, Azkona M, Latasa MU, García-Irigoyen O, Uriarte I, et al. Splicing regulator SLU7 preserves survival of hepatocellular carcinoma cells and other solid tumors via oncogenic miR-17-92 cluster expression. Oncogene. 2016;35(36):4719-29. Epub 2016/01/25. doi: 10.1038/onc.2015.517. PubMed PMID: 26804174.

93. Wang Y, Wang H, Li L, Li J, Pan T, Zhang D, et al. Elevated expression of STIM1 is involved in lung tumorigenesis. Oncotarget. 2016;7(52):86584-93. doi: 10.18632/oncotarget.13359. PubMed PMID: 27863410; PubMed Central PMCID: PMCPMC5349937.

94. Bhat R, Bhattacharyya PK, Ratech H. An Immunohistochemical Survey of SNARE Proteins Shows Distinct Patterns of Expression in Hematolymphoid Neoplasia. Am J Clin Pathol. 2016;145(5):604-16. doi: 10.1093/ajcp/aqw022. PubMed PMID: 27247366.

95. Carrozzo R, Dionisi-Vici C, Steuerwald U, Lucioli S, Deodato F, Di Giandomenico S, et al. SUCLA2 mutations are associated with mild methylmalonic aciduria, Leigh-like encephalomyopathy, dystonia and deafness. Brain. 2007;130(Pt 3):862-74. Epub 2007/02/14. doi: 10.1093/brain/awl389. PubMed PMID: 17301081.

96. Tu H, Ahearn TU, Daniel CR, Gonzalez-Feliciano AG, Seabrook ME, Bostick RM. Transforming growth factors and receptor as potential modifiable pre-neoplastic biomarkers of risk for colorectal neoplasms. Mol Carcinog. 2015;54(9):821-30. Epub 2014/04/09. doi: 10.1002/mc.22152. PubMed PMID: 24719252.

97. Rios R, Sangro B, Herrero I, Quiroga J, Prieto J. The role of thrombopoietin in the thrombocytopenia of patients with liver cirrhosis. Am J Gastroenterol. 2005;100(6):1311-6. doi: 10.1111/j.1572-0241.2005.41543.x. PubMed PMID: 15929762.

98. Meyer E, Michaelides M, Tee LJ, Robson AG, Rahman F, Pasha S, et al. Nonsense mutation in TMEM126A causing autosomal recessive optic atrophy and auditory neuropathy. Mol Vis. 2010;16:650-64. Epub 2010/04/13. PubMed PMID: 20405026; PubMed Central PMCID: PMCPMC2855733.

99. O'Hurley G, Busch C, Fagerberg L, Hallstrom BM, Stadler C, Tolf A, et al. Analysis of the Human Prostate-Specific Proteome Defined by Transcriptomics and Antibody-Based Profiling Identifies TMEM79 and ACOXL as Two Putative, Diagnostic Markers in Prostate Cancer. PLoS One. 2015;10(8):e0133449. doi: 10.1371/journal.pone.0133449. PubMed PMID: 26237329; PubMed Central PMCID: PMC4523174.

100. Liu Y, Jiang XL, Jiang DS, Zhang Y, Zhang R, Chen Y, et al. Toll-interacting protein (Tollip) negatively regulates pressure overload-induced ventricular hypertrophy in mice. Cardiovasc

bioRxiv preprint doi: https://doi.org/10.1101/158022; this version posted February 23, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Res. 2014;101(1):87-96. Epub 2013/11/26. doi: 10.1093/cvr/cvt232. PubMed PMID: 24285748; PubMed Central PMCID: PMCPMC3968303.

101. Feng Z, Prentice R, Srivastava S. Research issues and strategies for genomic and proteomic biomarker discovery and validation: a statistical perspective. Pharmacogenomics. 2004;5(6):709-19. doi: 10.1517/14622416.5.6.709. PubMed PMID: 15335291.

102. Dias MS, Hamel CP, Meunier I, Varin J, Blanchard S, Boyard F, et al. Novel splice-site mutation in TTLL5 causes cone dystrophy in a consanguineous family. Mol Vis. 2017;23:131-9. Epub 2017/03/18. PubMed PMID: 28356705; PubMed Central PMCID: PMCPMC5360453.

103. Ma Q, Li Y, Guo H, Li C, Chen J, Luo M, et al. A Novel Missense Mutation in USP26 Gene Is Associated With Nonobstructive Azoospermia. Reprod Sci. 2016;23(10):1434-41. Epub 2016/04/18. doi: 10.1177/1933719116641758. PubMed PMID: 27089915.

104. Gao Y, Wang Z, Hao Q, Li W, Xu Y, Zhang J, et al. Loss of ERα induces amoeboid-like migration of breast cancer cells by downregulating vinculin. Nat Commun. 2017;8:14483. Epub 2017/03/07. doi: 10.1038/ncomms14483. PubMed PMID: 28266545; PubMed Central PMCID: PMCPMC5344302.

105. Abdul S, Boender J, Malfliet JJ, Eikenboom J, Fijn van Draat K, Mauser-Bunschoten EP, et al. Plasma levels of plasminogen activator inhibitor-1 and bleeding phenotype in patients with von Willebrand disease. Haemophilia. 2017. Epub 2017/03/17. doi: 10.1111/hae.13206. PubMed PMID: 28306198.

106. Armaiz-Pena GN, Gonzalez-Villasana V, Nagaraja AS, Rodriguez-Aguayo C, Sadaoui NC, Stone RL, et al. Adrenergic regulation of monocyte chemotactic protein 1 leads to enhanced macrophage recruitment and ovarian carcinoma growth. Oncotarget. 2015;6(6):4266-73. doi: 10.18632/oncotarget.2887. PubMed PMID: 25738355; PubMed Central PMCID: PMCPMC4414188.

107. Wang S, Reeves B, Pawlinski R. Astrocyte tissue factor controls CNS hemostasis and autoimmune inflammation. Thromb Res. 2016;141 Suppl 2:S65-7. doi: 10.1016/S0049-3848(16)30369-3. PubMed PMID: 27207429.

108. Di Paolo NC, Shayakhmetov DM. Interleukin 1α and the inflammatory process. Nat Immunol. 2016;17(8):906-13. doi: 10.1038/ni.3503. PubMed PMID: 27434011; PubMed Central PMCID: PMCPMC5152572.